Syllable Structure, Frequency, Analogy, and Phonetics:
Factors in North Kyungsang Korean Accentuation of Novel Words

A Dissertation Presented

by

Hyun-ju Kim

to

The Graduate School

in Partial Fulfillment of the

Requirements

for the Degree of

Doctor of Philosophy

in

Linguistics

Stony Brook University

May 2012

Stony Brook University

The Graduate School

Hyun-ju Kim

We, the dissertation committee for the above candidate for the

Doctor of Philosophy degree, hereby recommend

acceptance of this dissertation.

Ellen I. Broselow – Dissertation Co-Advisor
Professor, Department of Linguistics

Marie K. Huffman – Dissertation Co-Advisor
Associate Professor, Department of Linguistics

Christina Y. Bethin – Chairperson of Defense
Professor, Department of Linguistics

Michael Kenstowicz
Professor, Department of Linguistics
Massachusetts Institute of Technology

This dissertation is accepted by the Graduate School

Charles Taber
Interim Dean of the Graduate School

Abstract of the Dissertation

Syllable Structure, Frequency, Analogy, and Phonetics:
Factors in North Kyungsang Korean Accentuation of Novel Words

by

Hyun-ju Kim

Doctor of Philosophy

in

Linguistics

Stony Brook University

2012

North Kyungsang Korean (NKK) is a pitch accent language in which each word has one of a restricted set of possible tonal patterns, and where the tonal pattern of a given lexical word is not fully predictable. This dissertation reports on a corpus study of accent patterns in existing words and the results of a study in which NKK speakers were asked to produce novel forms. This study demonstrated that when NKK speakers produce novel words, their accent patterns reveal regular tendencies, most notably a tendency for heavy syllables to attract accent. An experiment in which speakers were asked to produce novel forms that differed in only one segment from existing forms revealed that these tendencies do not originate from analogy to phonetically similar familiar words. Rather, they reflect a statistical association in the lexicon between accent and heavy syllables, though this association was even stronger in novel words than in existing words. In addition, phonetic factors predicting the position of accent were found: accent was more likely in syllables with aspirated onset consonants and in syllables containing high vowels, perhaps due to the acoustic property of higher F0 which is shared by high vowels, vowels following aspirated consonants, and accented (high-toned) vowels.

I argue that NKK speakers' behavior in accenting novel words reflects a set of universal markedness constraints. In native existing words, constraints which require lexical entries to surface faithfully in the output outrank these markedness constraints, but when no lexical entry is present, the effects of the markedness constraints emerge. I present a grammar involving a set of stochastically ranked constraints which predicts the patterns of both existing and novel words, and present evidence that this grammar is learnable on the basis of the patterns of existing words. The data from the accentuation of novel words supports the conclusion that speakers tend to

extend statistical tendencies of the lexicon to novel forms when such tendencies are consistent with cross-linguistic phonological tendencies.

Dedicated to my mother,

Tae-su Cho

TABLE OF CONTENTS

List of Figures

List of Tables

ACKNOWLEDGMENTS

I feel gratitude to the many people who helped me through the dissertation process. First, I would like to thank my two advisors Ellen Broselow and Marie Huffman. Ellen Broselow guided me at each step of my dissertation and assisted me in becoming an independent researcher in the long run. She inspired me with many insightful questions while writing my dissertation, which helped me think in a more scientific way and develop my idea as relevant to the field. Marie Huffman drew my attention to phonetics and refreshed my phonetician's perspective. I have no doubt that her comments and suggestions improved the quality of my dissertation research. Both of my advisors have been a blessing to my life and the best thing about my seven years at Stony Brook. Even after I left Stony Brook, I was able to continue my dissertation research in Korea thanks to their continuing support. Without their encouragement, it would certainly have been more difficult for me to complete my dissertation. I hope they know that they are my role models as scholars, teachers, and mentors.

Many thanks are also due to my committee members. I thank Christina Bethin for her many comments and suggestions. Christina Bethin was always supportive and meetings with her were always productive, and I have enjoyed discussing linguistics with her. My cordial appreciation goes to Michael Kenstowicz, who honored me by joining my committee and who inspired me with his insightful comments and helpful suggestions. I am truly grateful for his encouragement, which never let me doubt that this project was worthwhile.

I would like to thank all of the linguistics community at Stony Brook University. I have been lucky to have many outstanding teachers at Stony Brook and I learned a lot about linguistics from each of them. Particularly to be thanked are Dan Finer, who was my first QP advisor; Richard Larson, who was a great teacher, and who expressed care for my future directions after graduation; John Bailyn, whose syntax course was fun and interesting; Alice Harris, who guided her students to become strong field workers and real professionals; and Heejung Ko, who called my interest to language acquisition. I would also like to express gratitude to Bob Hoberman, who, although I missed out on having him as a teacher, was a wonderful professor to TA for.

I survived my life as a graduate student thanks to my fellow graduate students in the department and with the department assistant Sandra Brennan. Sandra Brennan took care of many administrative details throughout the years, as well as made our department friendlier and family-like. I would like to give special thanks to: Hijo Kang, who was always open to me to discuss many linguistic issues and made himself constantly available to help with whatever I asked for; Marlin Taylor, who was my delightful office-mate and great colleague on several projects; Andrei Antonenko, Mark Lindsay, and Yu-an Lu, who have helped me in innumerable ways and been my friends as well as colleagues; and not least, Jiwon Hwang, Miran Kim, and Youngran An for their special friendship and assistance at school and beyond. Also, I am grateful to Georgia Zellou, with whom I began the long journey as a linguist, and who has shown her everlasting affection and cheerful wishes for my academic success.

I am also deeply thankful for my professors in Korea: My cordial thanks go out to Professor Hyang-sook Sohn, who has been a great mentor in my academic career. At Kyungpook

National University (KNU), Prof. Hyang-sook Sohn inspired me to become a phonologist in the first place. She continued to show her support with encouragement while I was studying at Stony Brook. Professor No-ju Kim introduced me the interesting aspects of North Kyungsang Korean tone and helped me to extend my interests to theoretical linguistics. In addition, I am indebted to Professor Chung-hee Chung, who motivated me to study abroad and was a constant source of encouragement and care.

Much of this dissertation was based on experiments. Thanks to all NKK speakers who participated in the experiments. I especially thank Minji Sung and Yukyung Hong who served as consultants for NKK accent patterns in the corpus. In addition, Minji Sung provided her judgments on accent patterns in novel words, which took considerable time and efforts. I would like to express gratitude to Kwangmi Yu at Kyungpook National University, who has sent me timely encouraging messages that warmed my heart when times were difficult.

I can never express my thanks enough to my family, who has sustained me all the time. My deepest thanks and love go to my husband Iksu Youn and my sons San and Kang. Iksu Youn gave me constant love and encouragement, although he was the person who actually needed that. I have felt sorry for that he had to live a lonely life while I was off at school. San is a blessing to me in my life. He had been a great companion during my long journey at Stony Brook and brightened each day in the last 7 years of my life. Kang is the apple of my eye and refreshed me when I was weary from working on the dissertation. I also thank my brother and sisters, sisters-in-law, and brothers-in-law, all of whom supported me with their love. Finally, but certainly not the least, I would like to thank to my mother Taesu Cho. I would never have made it through without her aid and comfort throughout the years. There are no words that can express my deep gratitude to my mother, and I am dedicating this dissertation to her.

Chapter 1 Introduction

North Kyungsang Korean (henceforth, NKK) is a pitch accent language in which each word has one of a restricted set of possible tonal patterns, and where the tonal pattern of a given lexical word is not fully predictable. Thus, the position of the syllable bearing pitch accent must be specified in the word's lexical entry. However, I will present evidence that when lexical information for accent position is absent, some regular tendencies emerge in the accentuation of novel words. Specifically, I argue that NKK speakers' accentuation of novel words provides evidence for the presence of universal preferences for a connection between heavy syllables and high pitch. These tendencies have some support in the native language lexicon, where they are manifested as statistical preferences, but these preferences emerge even more strongly in the accentuation of novel words. The effects of these tendencies can be expressed as an effect of universal markedness constraints which are obscured in native vocabulary by higher-ranked faithfulness constraints, but which emerge when no lexical entry is present.

Although NKK accent patterns are not predictable for most native words, regular patterns can be found in at least some words: (i) penultimate accent is the default pattern for words longer than three syllables (Y-H Chung 1991; N-J Kim 1997); (ii) syllables containing a long vowel always have a high tone (Y-H Chung 1991; N-J Kim 1997). Accent patterns in loanwords also reveal some regular tendencies that differ from the patterns found in native vocabulary; e.g., CVC syllables tend to attract accent in loanwords (Kenstowicz and Sohn 2001). Kenstowicz and Sohn (2001) propose that the tendency for closed syllables to attract accent reflects the default UG setting which counts CVC syllables as heavy. I will show that statistical accent patterns in the native lexicon actually reflect these tendencies toward accented CVC syllables. Also, NKK speakers tend to accent novel words in accordance with these syllable structure-sensitive tendencies in the native lexicon, and in fact, these tendencies emerge even more strongly in the accentuation of novel words than in the patterns of existing words. This indicates that NKK speakers do not merely replicate the lexical frequency patterns, but rather exhibit a bias toward preferences for the association between accent and heavy syllables, resulting in stronger regular tendencies in the accent patterns of novel words.

In addition to syllable structure, phonetic factors such as aspiration and F0 influence accent placement in novel words. Syllables containing aspirated onset consonants are more likely to be associated with a high tone. Also, a high vowel tends to attract a high tone more often than a low vowel. These findings suggest that phonetic qualities also affect accent placement in novel words. Both the effects of aspiration and of vowel quality could be attributable to the connection of higher F0 and high tone: F0 of high vowels is intrinsically higher than F0 of low vowels (C-W Kim 1968; Ladd & Silverman 1984; Whalen and Levitt 1995; Yang 1996; Hoole *et al*. 2011); F0 of a vowel following an aspirated consonant is also higher than F0 of vowels following non-aspirated consonants (Han & Weitzman 1965; C-W Kim 1968; S-A Jun 1993; Silva 1992, 2006, M-R Kim 2000, M. Kim & Duanmu 2004, Kenstowicz & Park 2006). Since F0 is utilized for the tonal contrast in NKK, NKK speakers might perceive higher F0 as signaling accent.

A uniform model for both existing and novel words is proposed to account for the NKK accent patterns. This model adopts a set of stochastically-ranked constraints, which can predict the probabilistic outcomes that emerge in both existing and novel words. I argue that syllable structure-sensitive tendencies emerging in novel words are guided by a set of universal markedness constraints. In native existing words, constraints which require lexical entries to surface faithfully in the output outrank these markedness constraints, but when no lexical entry is present, the effects of the markedness constraints emerge. I present a stochastic grammar to account for both the native and novel word patterns, and demonstrate that this grammar is learnable on the basis of the lexical patterns in existing words.

1.1 Hypotheses

This dissertation begins with the broad question of what determines phonological behavior. It has been argued that a phonological grammar enables speakers to produce and understand words that they have never heard before. If so, patterns that emerge in the production of novel words should reflect an abstract internalized grammar. However, an alternative approach assumes that the production of novel words is mainly influenced by analogy with existing words. In this dissertation, I address this question in the case of accentuation of novel words in NKK, where accent is not fully predictable. How do NKK speakers accent novel forms when lexical entries are not available? Do they merely reproduce lexical patterns? Or is there a certain preference for a particular pattern?

The effects of analogy to patterns in existing words have been found in several different studies of stress in novel words (e.g. Baker and Smith 1976; Daelemans *et al.* 1994; Gillis *et al.* 2000; Eddington 2000; Guion *et al.* 2003; Face 2004). Guion *et al.* (2003) showed that the location of stress on phonologically similar real words influenced stress placement in English nonwords. Face (2004) also found that stress in phonetically similar words was one of the factors that influenced stress perception in Spanish nonwords. To account for analogical inference in linguistic behavior, many computational models have been proposed; e.g., exemplar-based models such as Generalized Context Model (GCM, model of similarity based classification: Nosofsky 1990); feature-based classification models (Tversky 1977) such as TiMBL (Daelemans *et al.* 2000); Analogical Modeling of Language (AML, Skousen 1989); and Minimal Generalization Learner (MGL; Albright and Hayes 2002). Most of these models suggest that an analogy-based model can generalize a certain linguistic pattern beyond the training data and produce patterns in a manner similar to that found in actual language learning. For example, Daelemans *et al.* (1994) argue that a learning algorithm based on similarity of new items to stored exemplars successfully predicts the patterns in novel words produced by human language learners.

In addition to the effect of word similarity, many recent studies have shown that lexical frequency affects linguistic behavior in different linguistic areas, including phonetics (Pierrehumbert 2002; Hay *et al*. 2004), language change (Bybee 1985, 2000, 2001; Bybee & Hopper 2001), and language acquisition (Jusczyk *et al.* 1994; Saffran 2003; Saffran & Thiessen 2003). For example, Hay *et al.* (2004) presented evidence that perceived well-formedness of consonant clusters in English nonsense words is influenced by lexical frequency of the clusters in the language. As for the role of frequency in sound change, Bybee (2000, 2001) demonstrates that a lenition process in English is related to word frequency in that more frequent words are more likely to undergo reduction. For example, /t/ deletion in double-marked past tense verbs (e.g., ‘left’, ‘felt’) is more likely in more frequent forms than in less frequent forms. Language acquisition data also provide additional evidence for frequency effects. For example, Jusczyk *et al.* (1994) found that nine-month-old infants prefer frequent phonotactic patterns in their language to infrequent ones.

Although classical generative phonology has assumed that frequency effects reside outside the grammar, recent research has presented evidence that lexical frequency interacts with grammatical effects (e.g. Anttila 1997; Boersma 1997; Boersma & Hayes 2001). For example, Frisch & Zawaydeh (2001) show that speakers of Jordanian Arabic use lexical statistical patterning in judging novel words with varying degrees of similarity between root consonants. Hay, Pierrehumbert, & Beckman (2004) also show that English speakers' well-formedness judgments of nonsense words containing nasal-obstruent clusters are gradiently related to the frequency of the clusters as well as a dispreference for two strident coronals within a word. Moreover, in infant learning of phonotactic patterns in an artificial language, Saffran & Thiessen (2003) showed that English nine-month-old infants used statistics in learning of the phonotactic patterns. However, the infants learned patterns that grouped all voiceless stops (e.g./p/, /t/, /k/) as a class of items comprising the first sound in artificial word tokens better than patterns that grouped mixed voicing combinations (e.g. /p/, /d/, /k/), although the mixed combinations were equally available in the statistical patterning. Additionally, Zuraw (2000) shows that Tagalog native speaker's judgments of novel words with the dispreference for a sequence of a nasal and a voiceless obstruent and for root-initial back nasals reflected the lexical frequency of such patterns in existing words.

In contrast, there have been studies that argue for a phonological grammar containing default rules which are uniformly applied when a lexical entry is not available. Shinohara (1997, 2000) found that despite the unpredictability of lexical accent patterns in Japanese native words, accent patterns in foreign words from French are fairly regular in that the antepenult is accented if the penult is light; otherwise, the penult is accented. Shinohara suggests that this pattern represents the default, which is also found in some native words such as certain compounds and proper names. Similarly, Crosswhite *et al.* (2003), in a study of how Russian speakers stress novel words which lack lexical specifications, found that the stem-final syllable was stressed in 80% of responses, even though the occurrence of stem-final stress in the lexicon is much less frequent (30%). The responses were surprisingly uniform with a high rate (70%) of unanimous responses, indicating that there was a clearly preferred position for stress. Therefore, Crosswhite *et al.* propose that there is a default stress location and the favored position is regulated by the phonology, independent of lexical frequency effects.

Given these findings, three hypotheses were investigated for NKK accentuation in novel words, as follows:

(1) *The Default Accent Hypothesis*: there is a default accent rule which applies when there is no lexical entry available.

Among the previous studies of NKK accent patterns, N-J. Kim (1997) actually proposes default penultimate accent for underlyingly accentless words containing no long vowel. Kim's hypothesis predicts penultimate accent as default regardless of syllable structure since native accentuation cannot be solely determined by syllable structure.

(2) *The Stochastic Accent Hypothesis*: accentuation in novel words will be consistent with statistical patterning in the lexicon.

Under this hypothesis, lexical frequency is related to accent patterns in novel words. Thus, if a certain preference for a particular pattern is found in novel words, the pattern should reflect patterns of lexical frequency.

(3) *The Word Similarity Effect Hypothesis*: accent patterns in novel words will follow the patterns of existing words that are phonetically similar to the novel words.

Under the analogy-based approach, in the absence of lexical accent specification, it is plausible that NKK speakers access neighboring lexical items that are similar to the novel words and follow the patterns of these similar words.

In addition to these three hypotheses, I consider one additional hypothesis connecting accent on novel forms to phonetic factors. Korean vowels following aspirated consonants have higher fundamental frequency (F0) than vowels following lax consonants. Since that higher F0 is a mark of lexical tonal contrasts in NKK, the effect of aspiration on F0 of a following vowel could be utilized when assigning accent in cases where no lexical entries are available. As for the phonetic correlates of the laryngeal contrast among lax, tense and aspirated stops in Korean, Voice Onset Time (VOT) as well as F0 has been considered as one of the features signaling the phonation type contrast in standard Korean and in Kyungsang Korean (Han & Weitzman 1970; Silva 1992, 2006; M-R Kim 2000; M. Kim and Duanmu 2004; Kenstowicz and Park 2006, and many other studies). However, Silva (2006) recently found that F0 rather than VOT is more likely to serve as a primary cue for the distinction between lax and aspirated stops because neutralization of the VOT distinction between lax and aspirated stops has been in progress for some younger speakers. Silva argued that contemporary Seoul Korean, which is not a pitch accent language, has developed a tonal system to encode the phonemic contrast between lax and aspirated stops: lax stops are associated with a low tone, while aspirated stops are marked with a high tone. Based on these previous findings, the effect of aspiration on NKK accentuation in novel words is hypothesized as follows:

(4) *The Phonetic Effect Hypothesis*: phonetic factors will influence NKK accentuation in novel words.

This hypothesis predicts that if NKK speakers are sensitive to the association between aspiration and higher F0, vowels after aspirated onsets will be more likely to attract accent in novel words.

This dissertation reports on several experimental studies which were performed using novel words to test these hypotheses, and to identify the factors determining accent placement when no lexical entry is available.

1.2 Dissertation Outline

The dissertation is organized as follows.

Chapter 2 provides the linguistic background on NKK lexical accent patterns based on previous research, addressing the phonological basis underlying the NKK lexical accent patterns within an Optimality Theory framework.

Chapter 3 presents an experimental study testing the *Default Accent Hypothesis*, which predicts that penultimate accent will be the most common pattern in novel words. Contra the predictions of this hypothesis, the results show that while penultimate accent was the most frequent pattern in words lacking CVC syllables, CVC syllables show a significant tendency to attract accent. These results suggest that the preferred pattern is for heavy syllables to attract accent, rather than for the penultimate syllable to be accented, regardless of the syllable structure of the word. This chapter also examines additional potential factors in accent patterns in novel words: effects of coda type (obstruent vs. sonorant) and effects of vowel quality (high vs. low). The results show no consistent effect of coda quality, though high vowels were more likely than low vowels to be associated with a high tone in novel forms.

Chapter 4 presents a corpus study examining the accent patterns of existing words, testing the *Stochastic Accent Hypothesis*, which predicts that accentuation in novel words will be consistent with statistical patterning in the lexicon. This study revealed that existing words do indeed exhibit a tendency for heavy syllables to be accented, although this tendency was weaker in existing than in novel words.

Chapter 5 presents an experimental study testing the *Word Similarity Effect Hypothesis*, which predicts that the accent patterns of novel words will follow those of existing words that are phonetically similar. Counter to the predictions of this hypothesis, the effects of analogy to phonetically similar words were not significant in accent patterns in novel words. This study found some cases where analogy to existing words was a related factor, but syllable structure-sensitive tendencies in novel words were never overridden by the analogy effects.

Chapter 6 presents a formalized account of the accent patterns in novel words based on stochastic constraint rankings. This chapter proposes that a stochastic grammar, constructed on the basis of native lexical patterns (e.g. Zuraw 2000; Hayes & Londe 2006), provides an explanation of the syllable structure-sensitive tendencies emerging in the novel words. Also, this chapter suggests that NKK speakers learn this grammar via exposure to the lexicon. A simulation assuming the Gradual Learning Algorithm (GLA)(Boersma1997; Boersma & Hayes 2001) demonstrates that the proposed stochastic grammar was learnable given the training data from the corpus.

Chapter 7 presents an experimental study performed to test the *Phonetic Effect Hypothesis*, which predicts that the accent patterns of novel words will be sensitive to onset quality (aspirated vs. lax). This chapter provides evidence that aspiration plays a role in the assignment of a pitch accent pattern to novel words. The results show that penultimate accent, which was most frequent in a word without aspirated consonants, was less likely when a word contains aspirated onsets. In contrast, double accent was more common in trisyllabic words whose initial syllable has aspirated onsets and final accent was assigned more often in bisyllabic words with aspiration on final onsets. This suggests that phonetic effects were also important as a factor in explaining the variable accent patterns of newly adopted words.

Chapter 8 outlines the conclusions of the dissertation: that syllable structure plays an important role in accent placement in novel words, and that the syllable structure-sensitive patterning is consistent with the patterns found in the lexicon, although the preference for association of accent with heavy syllables emerges significantly more strongly in novel than in existing words. Analogy to patterns of phonetically similar words is not significantly involved in accent placement in novel words, but interacts cumulatively with syllable structure. Finally, phonetic effects, in the form of a tendency for syllables with aspirated onsets to attract accent, appear as a relevant factor, serving as a source of variation in accent placement. These findings suggest that the accent patterns emerging in novel words are phonologically-driven and attributable to hidden rankings among markedness constraints, which are inactive but present covertly in the native phonology.

Chapter 2 Background: Lexical Accent Patterns in North Kyungsang Korean

2.1 Introduction

In North Kyungsang Korean (henceforth, NKK), each word has a lexically specified tone pattern, as illustrated in the following minimal triple *káci* 'kind', *kacî* 'eggplant', *kácî* 'branch'. Nevertheless, NKK accent placement is not completely random, since not all logically possible accent patterns occur in NKK.

Some scholars have classified Kyungsang Korean as a tone language (e.g. W. Huh 1965, 1985; C-G Gim 1977, 1985; Lee 1994,1997), arguing for three distinctive tones, High, Mid, and Low, based on the three distinct pitch levels. Under the tone language approach, low tone and mid tone are contrastive and each syllable in a word is assigned one of these three tones. An alternative and more widely accepted analysis is that NKK is a pitch accent language, in which a particular syllable can be lexically specified as accented (e.g. Narahara 1985; G-R Kim 1988; J-W Kim 1991; N-J Kim 1997; Kenstowicz & Sohn 2001; Jun *et al*. 2006; D-M Lee 2008a, 2008b). An accented syllable is always high, and a syllable following this accented syllable within a word has a low tone. In accordance with this view, J-W Kim (1991) explicitly argued that at the phonological level, only a two-way tonal contrast is necessary, with three different pitch levels resulting from the presence of underlying accent. Namely, a high tone can be either: underlyingly accented or underlyingly unaccented. Therefore, it is possible that NKK speakers may perceive an accented high tone as higher than an unaccented high tone. D-M Lee (2008b) characterizes the tonal patterns of NKK words as follows: i) the rightmost high tone is followed by a low tone; ii) a prosodic word contains no more than one pitch fall; iii) a sequence of three high tones cannot appear within a prosodic word; iv) a NKK word (unlike South Kyungsang Korean words) may begin with a sequence of two low tones.

The lexical accent patterns in NKK have been analyzed within different theoretic frameworks; e.g. autosegmental theory (G-R Kim 1988; Y-H Chung 1991); Optimality Theory (N-J Kim 1997); and metrical grid theory (S-H Kim 1999). However, all these studies agree that the accent pattern of NKK words is at least partially lexically determined. Therefore, NKK speakers must consult lexical information when it comes to accentuation in a lexical word, but the grammar of NKK defines only a subset of possible tonal patterns as grammatical.

2.2 NKK Accent Patterns

2.2.1 Possible patterns vs. Impossible patterns

The lexical accent patterns in NKK can be grouped into four patterns: initial, final, penultimate and double accent:

(1) Lexical accent patterns in NKK[1]

a. Initial accent

kə́puki	'turtle'	kámani	'straw rice-bag'

b. Final accent

cintallé	'azalea'	satarí	'ladder'

c. Penultimate accent

apə́ci	'father'	camcári	'dragon fly'

d. Double accent

kícíke	'a stretch'	múcíke	'rainbow'

Double accent (1d) is restricted to the initial two syllables. Furthermore, monomorphemic words longer than three syllables overwhelmingly take accent on the penultimate syllable:

(2) Penultimate accent in longer words

kosɨmtóc^{h}i	'porcupine'
aciráŋi	'haze'

G-R Kim (1988) reports that initial accent (HLLL) and double accent (HHLL) are also possible in quadrisyllabic monomorphemic words, while N-J Kim (1997) and Y-H Chung (1991) argue that only penultimate accent (LLHL) is possible (cited by D-M Lee 2008b). I have found that quadrisyllabic words with non-penultimate accent are very rare in the lexicon. Furthermore, penultimate accent is also possible alternatively for the words which are argued to carry non-penultimate accent; e.g. *hárápəci* vs. *harapə́ci* 'grandfather'; *ə́rápəni* vs. *orapə́ni* 'brother'. Therefore, I assume that penultimate accent is a typical pattern for words longer than three syllables.

In summary, the attested accent patterns, where H indicates high tone and L indicates low tone, are (a) penultimate accent in words longer than three syllables(LLHL); (b) double accent on the initial two syllables of bisyllabic or trisyllabic words (HH; HHL); (c) initial, penultimate, or final accent in words of one syllable to three syllables (HLL; LHL; LLH). The next section will discuss the nature of the grammar that generates possible patterns like HLL, LLH, LHL, HHL but bans unattested patterns like LLL, LHH, HLH, HHH.

2.2.2 Constraints

First of all, the accent pattern of a specific lexical item is not necessarily predictable. Hence, it has been assumed that an underlying representation (UR) must contain the lexical information for accent patterns for at least some words. For example, the accent pattern for specific words such as *kə́puki* 'turtle', *satarí* 'ladder', and *apə́ci* 'father' depends on information that must be contained in the input. In accordance with the principle of Lexicon Optimization proposed by Prince & Smolensky (1993: 191), whereby the most faithful underlying-surface pair is the optimal pair, NKK speakers would assume that the position of underlying accent is the same as

[1] In the transcriptions, *ə* stands for the mid unrounded vowel, *ɨ* for the high central/back unrounded vowel, *c* for an alveolar affricate, and *c^h* for an aspirated alveolar affricate.

appears in the surface form. Thus, for a word class with initial accent (e.g. *kə́puki* ‘turtle’), the input should be a form with a high tone on the initial syllable (|σ́σσ|), whereas for a word class with final accent (e.g. *satarí* ‘ladder’), the final syllable should have a high tone in the input (|σσσ́|). The faithful constraint IDENT(T) ensures faithful mapping of a high tone from the input to the output.

(3) IDENT (T) (de Lacy 2002):
If mora *x* bears tone T in the input, then the output correspondent of *x* bears T.

However, accent patterns like LLL, LHH, HLH, HHH never appear in surface forms in NKK. According to the Richness of the Base hypothesis, underlying representations should not be restricted, and thus any underlying pattern should be available in the input. Therefore, markedness constraints ranked higher than the faithfulness constraint IDENT(T) must rule out these patterns even though these unattested forms are possible in the input. Because at least one syllable in every NKK prosodic word carries high tone on the surface form, I assume the following constraint (4), which rules out LLL forms:

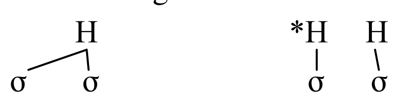
(4) ..H..]$_{PW}$: A prosodic word must have a high tone (N-J Kim 1997; D-M Lee 2008b)[2]

Only one pitch peak is present in a prosodic word in NKK. The constraint OCP(H) (N-J Kim 1997; Y-H Chung 2006) prohibits patterns like HLH, LHLH:

(5) OCP (H): More than one high tone is not allowed within a domain (prosodic word)

In accordance with OCP (H), words with double high tones in (1) are analyzed as containing a single high tone which is associated with two syllables, as shown in (6).

(6) Tone sharing



Since double high tones are possible only in word-initial position, ALIGN-L (DOUBLE) is suggested to prohibit LHH forms, since double accent always occurs word-initially not word-finally (HHL vs. *LHH).

(7) ALIGN-L (DOUBLE): Align the left edge of a shared high tone with the left edge of a word

Finally, HHH forms never appear in NKK. To ban HHH forms, the constraint HLINK should be active in NKK, which explicitly prohibits a high tone from spreading over more than two syllables.

[2] The constraint H]$_{PW}$ is comparable to *Toneless: high toneless words are prohibited (N-J Kim 1997), and One-H-Pw : Only one high tone is present in a prosodic word (D-M Lee 2008b).

(8) HLINK: A high tone may not be linked to more than two tone bearing units (Zec 1999, cited by Y-H Chung 2006, 10)

Assuming that these constraints $H]_{PW}$, OCP(H), ALIGN-L (DOUBLE) and HLINK are undominated, only four patterns (initial accent, penultimate accent, final accent, and double accent on the initial two syllables) are grammatical, since these constraints ban the unattested patterns (HHH; LHH; HLH; LLL). The following tableau illustrates the evaluation of these constraints, showing that patterns like LLL, LHH, HLH, and HHH will never surface.

(9) Attested patterns vs. Unattested patterns

Surface form	$H]_{PW}$	OCP	ALIGN-L(DOUBLE)	HLINK	IDENT(T)
a. HLL					
b. HHL					
c. LLH					
d. LHL					
☜e. LHH			*!		
☜f. LLL	*!				
☜g. HHH				*!	
☜h. HLH		*!			

In the evaluation (9), unattested patterns such as LHH(9e), LLL(9f), HHH(9g), HLH[3](9h), indicated by the symbol ☜, are ruled out due to the crucial violation of one of the undominated constraints, although they all satisfy IDENT(T). In contrast, attested patterns like HLL(9a), HHL(9b), LLH(9c), LHL(9d) all survive.

The following tableau illustrates the role of the faithful constraint IDENT(T), which ensures faithful realization of underlying accent on the surface form. I assume that the lexical pitch accent is represented by a high tone in UR, similar to Narahara's concept of 'prelinking' of tones (1985) (e.g. G-R Kim 1988; Y-H Chung 1991; N-J Kim 1997).

(10) Examples: *kə́puki* 'turtle' with initial accent; *apə́ci* 'father' with penultimate accent

/ kə́puki/	$H]_{PW}$	OCP	ALIGN-L(DOUBLE)	HLINK	IDENT(T)
☞ a. HLL					
b. HHL					*!
c. LLH					*!
/apə́ci/	$H]_{PW}$	OCP	ALIGN-L(DOUBLE)	HLINK	IDENT(T)
a. HHL					*!
b. LLH					*!
☞c. LHL					

[3] Candidate (9h) is assumed to have two high tones in output since spreading is blocked due to NO CROSSING ASSOCIATION LINE constraint. In terms of tone splitting, we need to posit the constraint INTEGRITY(T)(McCarthy & Prince 1995) undominated. Thus, (9h) cannot be a winning candidate.

INTEGRITY(T): Every input has only one output correspondent.

As shown in (10), words with initial accent are assumed to have a high tone on the initial syllable in UR, while words with penultimate accent in the surface form are assumed to have a high tone on the penultimate syllable in UR. Due to the constraint IDENT (T) which obliges the underlying accent to be realized faithfully on the surface form, initial accent is chosen for *kəpuki*, while penultimate accent is assigned in a word *apəci* which has underlying accent on the penultimate syllable.

2.2.3 Accentless group and default accent

Several previous studies actually postulate that there is a word group for which no underlying accent is present, which will be called an *accentless* group in this chapter (i.e. Y-H Chung 1991; N-J Kim 1997; S-H Kim 1999). However, the decision on which pattern is assigned in the accentless group is controversial. Y-H Chung (1991) and S-H Kim (1999) assume that the default position is actually final, based on the fact that the accent in a final accented stem shifts to a suffix:

(11) *Accent patterns in a stem with a suffix longer than monosyllabic*

camcári 'dragon fly' + cóc^{h}a 'even' → camcári coc^{h}a

múcíke 'rainbow' + cóc^{h}a 'even' → múcíke coc^{h}a

satarí 'ladder' + cóc^{h}a 'even' → satari cóc^{h}a

As shown in (11), final accent is less stable than other accent patterns since suffixes longer than monosyllabic will take the accent that is otherwise realized on the final syllable of the stem, as in *satari-cóc^{h}a* 'even ladder'. However, final accent is one of the least frequent patterns in native words longer than bisyllabic (e.g. 6% of trisyllabic lexical words) and uncommon in loanwords unless words contain heavy syllables word-finally (see Chapter 4 for the lexical statistics). Furthermore, the following chapter will show that final accent was the least favored pattern in novel words, which seems contrary to the predictions of default accent on novel words without underlying accent.

Alternatively, N-J Kim (1997) argues that words with penultimate accent belong to the accentless word group, in which penultimate accent is derived by a default phonological rule. N-J Kim argues that overwhelming penultimate accentuation in quadrisyllabic words supports the default penultimate accent, assuming that quadrisyllabic words lack underlying accent. N-J Kim (1997) proposes the following constraints in (12) for default accent assignment of words without an underlying high tone. The constraint NONFINALITY should be ranked higher than ALIGN-R(H) since penultimate accent wins over final accent in quadrisyllabic words.

(12) ALIGN-R (H): Align the left edge of a high tone with the right edge of a word.

NONFINALITY: Alignment of a high tone must not be final in a prosodic domain.

I assume that the constraints NONFINALITY and ALIGN-R (H) are violable in NKK because initial/penultimate/final accent patterns are all possible in words containing less than four syllables. Thus, it is assumed that these constraints should be ranked lower than the undominated constraints such as HLINK and OCP. In accordance with the assumption that quadrisyllabic words have no underlying high tone, the evaluation of the constraints is illustrated in the following tableau (13).

(13) Penultimate accent in quadrisyllabic words

	H]$_{PW}$	OCP	ALIGN-L(DOUBLE)	HLINK	IDENT(T)	NON-FINAL	ALIGN-R
a. LLLL	*!						
b. HHHH				*!			
c. LLHH			*!			*	
d. LHLH		*!				*	
e. HLLL							**!*
f. HHLL							**!
g. LLLH						*!	
☞h. LLHL							*

As shown in (13), penultimate accent (13h) is the best choice based on the evaluation of these markedness constraints. However, this analysis holds only under the assumption that no underlying high tone is present in this word.

2.3 Syllable Weight Sensitivity in NKK Accent Patterns

It has been widely observed that syllable weight plays an important role in certain phonological processes in many languages (e.g. McCarthy 1979; Hyman 1985; Hayes 1989, 1995; Broselow *et al*. 1997; Gordon 1999, 2004, etc.). In syllable weight-sensitive phenomena, syllables containing a long vowel (CVV) are usually treated as heavy in some languages: for example, CVV syllables always attract stress in Khalkha Mongolian (Walker 1995). In addition, CVC syllables are often treated as heavy in many languages. If CVC syllables are considered as heavy in a language, then CVV syllables are also heavy, but not vice versa. For example, in Finnish stress patterns, CVC syllables pattern with CVV syllables in that both CVC and CVV syllables always attract stress (Sadeniemi 1949, cited by Gordon 2004).

Previous studies of the NKK accent system have actually reported that syllable weight is also an important factor which influences NKK lexical accent patterns: syllables with a long vowel are always aligned with a high tone (e.g. *maŋné:* 'the youngest', *kó:kúma* 'sweet potato') (Y-H Chung 1991; N-J Kim 1997). As we see in (14), the final heavy syllable always attracts a high tone. The following examples are adopted from N-J Kim (1997).[4]

[4] A long vowel always occurs at the edges of the stem (Y-H Chung 1991). However, long vowels rarely occur word-finally: i) only a few bisyllabic words with a final heavy syllable are available: the examples in (14) are all that are known; ii) trisyllabic words (or words longer than trisyllabic) where a final syllable contains a long vowel are not attested at all.

(14)	apə́:m	'father'	kó:kúma	'sweet potato'
	əmə́:m	'mother'	kó:mpháŋi	'fungus'
	halpə́:m	'grandfather'	hó:ráŋi	'tiger'
	maŋné:	'the last child'		

However, CVC syllables do not pattern with CVV syllables in that CVC syllables do not necessarily attract a high tone pitch accent (e.g. *cin.tal.lé.*'azalea', *cam.cá.ri.*'dragon fly', *ə́.rí.sin.*'older person'). Therefore, it is apparent that CVC syllables are not counted as heavy in NKK native accent patterns (Kenstowicz & Sohn 2001).

The predictability of the weight-sensitive patterning suggests that the constraint WEIGHT-TO-TONE PRINCIPLE (WTP) (N-J Kim1997; Y-H Chung 2006; D-M Lee 2008b), which forces heavy syllables to have a high tone, should outrank the constraint IDENT (T). Thus, even if there is a heavy syllable without a high tone in the input, the heavy syllable should be realized with a high tone in the output due to the high-ranked constraint WTP.

(15) WEIGHT-TO-TONE PRINCIPLE (WTP): Heavy syllables must have a high tone[5]

As shown in (16), the OT evaluation for an example word *maŋné:* 'the last child' selects the form with a high tone on a long vowel in the surface form, regardless of the underlying accent position.

(16) Weight-sensitive accent assignment

(i) Input: /máŋne:/	WTP	IDENT(T)	NONFINALITY	ALIGN-R
a. máŋne:	*!			*
☞b. maŋné:		*	*	
(ii) Input: /maŋné:/	WTP	IDENT(T)	NONFINALITY	ALIGN-R
a. máŋne:	*!	*		*
☞b. maŋné:			*	

Whether we suppose a word-initial high tone (16i) or a word-final high tone (16ii) for the input of a word *maŋne:*, the word ends up with a high tone on the final heavy syllable.

When an initial syllable is heavy, double accent is always assigned over the initial two syllables, as in the following examples: *kó:kúma* 'sweet potato', *kó:mpháŋi* 'fungus', *hó:ráŋi* 'tiger'. N-J Kim (1997) analyzed these words having a floating high tone, which should surface as doubly linked to two syllables. This assumption applies to all double-accented words, whether they contain a long vowel or not (e.g. *múcíke* 'rainbow' vs. *kó:kúma* 'sweet potato'). However, I assume instead that the group of words containing a long vowel is assigned double accent due to the constraint WTP: syllables containing a long vowel always have a high tone. However, it is not clear what makes double accent preferable to initial accent, since initial accent should also

[5] WEIGHT-TO-STRESS PRINCIPLE proposed in Prince and Smolensky (1993) is revised to WEIGHT-TO-TONE PRINCIPLE for tone assignment.

satisfy WTP. I suggest that the interaction of WTP and ALIGN-R accounts for the assignment of double accent, which will be illustrated in the following example.

As shown in (17), double accent is chosen for a word *kó:kúma* 'sweet potato', regardless of the underlying position of a high tone. In terms of underlying accent position of the word *kó:kúma*, three possibilities are considered: (i) initial accent; (ii) final accent; (iii) penultimate accent. As illustrated in the tableau (17), the surface accent position is not necessarily faithful to the underlying position due to the effect of WTP.

(17) Possible inputs: /kó:kuma/ vs. /ko:kumá/ vs. /ko:kúma/

(i) Input: /kó:kuma/	WTP	IDENT(T)	NONFINALITY	ALIGN-R
a. kó:kuma				*!*
b. ko:kúma	*!	*		*
c. ko:kumá	*!	*	*	
☞d. kó:kúma				*
(ii) Input: /ko:kumá/	WTP	IDENT(T)	NONFINALITY	ALIGN-R
a. kó:kuma		*		*!*
b. ko:kúma	*!	*		*
c. ko:kumá	*!		*	
☞d. kó:kúma		*		*
(iii) Input: /ko:kúma/	WTP	IDENT(T)	NONFINALITY	ALIGN-R
a. kó:kuma		*		*!*
b. ko:kúma	*!			*
c. ko:kumá	*!	*	*	
☞d. kó:kúma				*

The output with double accent is chosen regardless of the underlying accent position. In the case (17i) when the input contains a high tone on the initial syllable, the faithful candidate with initial accent (17ia) fails due to the crucial violation of ALIGN-R. When the input contains a high tone on a non-heavy syllable as in (17ii) and (17iii), the faithful candidates (17iic) and (17iiib) are also ruled out by WTP. Therefore, regardless of the underlying position, a high tone falls on the initial heavy syllable, but spread to the following syllable when words contain a heavy syllable word-initially. Tone spreading is motivated by ALIGN-R, which makes a crucial role to select double accent over initial accent in a word beginning with a heavy syllable. Additionally, the higher-ranked constraints ALIGN-L(DOUBLE) and HLINK, which restrict double accent to word-initial position and prevent association of a high tone with more than two syllables, ensure the assignment of double accent in this word.[6]

[6] D-M Lee (2008a) performed a phonetic study of North Kyungsang and South Kyungsang accent and found that a pitch peak in words with double accent is realized on the second high tone. Given this fact, D-M Lee suggested that the rightmost high tone is underlyingly accented in double accented forms, whereas the initial high tone is

I have shown that double accent rather than initial accent is assigned in the cases where an initial syllable contains a long vowel (CVV.CV.CV), although the initial accent is also possible in the native accent patterns. This raises a question of how to explain the availability of initial accent in a word containing no long vowel (e.g *kə́puki* 'turtle'). In other words, why should underlying initial accent surface with spreading in words containing a long vowel word-initially, though the underlying initial accent is realized without spreading in words without a long vowel (e.g. *kə́puki*)? However, a corpus study reveals that the initial accent pattern in trisyllabic words is extremely rare: only 10 out of 411 trisyllabic words (2%) have the initial accent pattern (see Chapter 4 for more details about the lexical statistics). Furthermore, most of these words allow penultimate accent alternatively; e.g. *kə́puki* 'turtle'≈ *kəpúki*; *kámani* 'straw rice-bag' ≈ *kamáni*; *kámulchi* 'mullet'≈ *kamúlchi*. The variation in the initial accent word group contrasts with the final accent group, which also was a minor pattern in terms of the lexical frequency (7%: 29/411) but which rarely exhibited alternative patterns. I argue that the rarity of an initial accent pattern in the lexicon and the presence of alternative patterning for the initial accent are due to the effect of the constraint ALIGN-R, although its effect is not explicit in the lexical accent patterns.

On the other hand, the phonemic status of vowel length has been controversial. Although traditional analyses have assumed that Korean has a phonemic vowel length contrast (e.g. C-H Lee 1955; M-S Han 1964; Crothers 1978); e.g. *nun* 'eye' vs. *nu:n* 'snow', the phonemic distinction is not necessarily maintained in the younger generation of native speakers of standard Korean (e.g. J-K Park 1985; Magen and Blumstein 1993). Magen and Blumstein (1993) found that the duration of historically long vs. short vowels was not consistent across speech rates, suggesting that the role of vowel length as a phonemic contrast is being lost by younger Korean speakers. As for NKK speakers, I assume that at least younger NKK speakers are also losing the contrast as well. C-G Gim (1994) found that the accent patterns in a word group which contains a long vowel were distinct in NKK speakers in their sixties vs. in their twenties. C-G Gim describes that the pattern of double accent in words containing a long vowel is distinct from that of double accent in words with no long vowels: a long tone (historically developed from a rising (LH) tone in Middle Korean); e.g. *hó:ráŋi* 'tiger' vs. a short tone (historically developed from a high tone in Middle Korean); e.g. *múcíke* 'rainbow'. However, the contrast of tone length between the two double-accented groups was not maintained in the younger dialect, while the tone distinction was preserved in the older dialect: the younger NKK speakers produced both word groups (words with a long vowel; words with no long vowel) with a short tone pattern unlike the older NKK speakers. I speculate that the diverse patterns between the two age groups should be attributable to the different production of long vowels: probably the long vowels produced by younger NKK speakers are produced with shorter duration than the corresponding long vowels produced by the older NKK speakers. However, no systematic phonetic research on the production of long vowels between different age groups of NKK speakers has been found yet. We leave this issue for future research. This study of accent production in novel forms reported on in this dissertation employed novel words containing only short vowels.

---

underlyingly unaccented. D-M Lee's findings are in accordance with Michael Kenstowicz's comments: phonetic studies of South Kyungsang Korean have shown that the double accent is mid-tone on first syllable and then higher on second syllable. Regardless of the location of underlying accent, the current system predicts the same output.

2.4 Coda Consonant Effects in Loanwords

2.4.1 Syllable weight in loanwords

NKK accent patterns in loanwords reveal some regular tendencies that differ from the patterns found in native vocabulary (Kenstowicz and Sohn 2001; Y-H Chung 2002, 2006). Recall that coda consonants do not contribute to syllable weight in NKK native vocabulary. However, CVC syllables tend to attract accent in loanwords, as shown in (18).

(18) Accented CVC syllables in loanwords

kawún	'gown'	keím	'game'
sɨt^{h}ak^{h}íŋ	'stocking'	pɨracíl	'Brazil'
k^{h}élkəri	‘Calgary’	símp^{h}óni	'symphony'[7]

When a word contains no CVC syllables, penultimate syllables tend to be accented:

(19) Penultimate accent without CVC syllables

raít^{h}a[8]	'lighter'
ratío	'radio'

The preference for accented CVC syllables in loanwords, shown in (18), does not come from NKK native phonology since CVC syllables do not necessarily attract accent in native words (e.g. *cintallé* ‘azalea’; *camcári* ‘dragon fly’). Furthermore, as the patterns in (18) and (19) indicate, English stress is not always preserved in borrowed words.

Kenstowicz and Sohn (2001) propose that the tendency for closed syllables to attract accent reflects the default UG setting which counts CVC syllables as heavy, evidence of an “emergence of the unmarked” effect (McCarthy & Prince 1994) of the type that has been argued to emerge in loanword and second language phonology (Shinohara 1997, 2000; Broselow, Chen and Wang 1998, among others). Y-H Chung (2006) pointed out that CVC syllables count as heavy only word-initially or word-finally but not in word-medial position in NKK loanwords (e.g. *sɨt^{h}entátɨ* ‘standard’; *sɨt^{h}enp^{h}ótɨ* ‘Stanford’) and argued that these patterns are accounted for by the OT constraints, WTP, MORAICITY, and EDGEMOST:

(20) MORAICITY: From the nucleus on, each segment of a syllable is moraic. (Prince and Smolensky 1993, cited by Y-H Chung 2006, 8)

EDGEMOST (Heavy:L|R; Wd): A heavy syllable lies at either the left or the right edge of the word. (Prince and Smolensky 1993, cited by Y-H Chung 2006, 8)

MORAICITY requires a coda consonant to earn a mora. When EDGEMOST dominates MORAICITY, coda consonants of word-medial CVC syllables should not count a mora, while coda consonants

[7] When the initial heavy syllable attracts accent, it is realized as double accent as in *k^{h}élkʌ́ri* ‘Calgary’; *símp^{h}óni* ‘symphony’, which is consistent with the patterns in native words containing a long vowel as in *kó:kúma* ‘sweet potato’; *kó:mp^{h}áŋi* ‘fungus’.

[8] The vowel sequences *a.i* in *ra.i.ta.* and *i.o* in *ra.ti.o.* are not diphthongs in Korean and each vowel is the nucleus of a syllable.

of word-initial and word-final CVC syllables are rendered a mora. WTP prohibits a heavy syllable with a low tone. Thus, the prosodic position-sensitive coda moraicity emerging in loanwords is explained based on the interaction of these constraints. Y-H Chung explained the discrepancy of syllable weight between loanwords and native words with reranking of the constraints, proposing that these constraints are ranked low in the native phonology, since no such effects are present in native words, but promoted in loanword phonology.

2.4.2 Accent variation in loanwords

When a loanword contains more than one CVC syllable, different accent patterns are possible:

(21) Variation

séntál	'sandal'	piltíŋ	'building'
k^{h}épt^{h}ín	'captain'	nepk^{h}ín	'napkin'
t^{h}émpə́rin	'tambourine'	k^{h}ontɨról	'control'
íntə́net	'internet'	alp^{h}apét	'alphabet'

In bisyllabic words where both syllables are closed, final accent as well as double accent is attested, as in the following examples: *séntál* 'sandal'; *piltíŋ* 'building'. In trisyllabic loans which contain CVC syllables at the both edges (word-initially and word-finally), double accent as well as final accent is available; e.g. *t^{h}émpə́rin* 'tambourine'; *k^{h}ontɨról* 'control'. Kenstowicz and Sohn (2001) found that when two heavy syllables compete with a sonorant coda vs. with an obstruent coda, the syllable with a sonorant coda is more likely to attract accent than the syllable with an obstruent coda (e.g. *íntə́net* 'internet'; *mémpə́ʃip* 'membership'). Furthermore, an initial syllable closed by a geminate lateral tends not to carry accent (e.g. *sollomón* 'Solomon'). D-M Lee (2008b) also observed that CVC syllables closed by a geminate lateral resist a high tone in South Kyungsang Korean (SKK) loanword patterns. D-M Lee suggests that an initial syllable has priority in attracting a high tone in heavy-light-heavy SKK loanwords: 69% (20/29) of the trisyllabic loans with the heavy-light-heavy combination showed a high tone on the word-initial heavy syllable, after excluding final accented heavy-light-heavy words with an initial heavy syllable closed by a geminate lateral.

Nevertheless, exceptions to these generalizations exist. First, in some words, a syllable with an obstruent coda wins over a syllable with a sonorant coda, as in the examples *alp^{h}apét* 'alphabet', *intənét*[9] 'internet'. Second, there are cases in which the initial syllable fails to attract accent even where the initial heavy syllables are closed by non-geminate lateral or non-lateral codas, as in the examples *piltíŋ* 'building', *alp^{h}apét* 'alphabet', *k^{h}ontɨról* 'control'. Furthermore, the preference for word-initial/final accented heavy syllables is contradicted by the following examples *ollímpik* 'olympic'; *t^{h}əkʃíto* 'tuxedo'. The presence of variation and exceptional cases contra the major tendencies emerging in loanwords implies that accent patterns emerging in

[9] Both double accent and final accent is possible for a word *intənet* 'internet' due to NKK speakers' variation. I observed that both patterns are fairly common.

loanwords cannot be described by a strictly categorical grammar, though the majority of the loanwords follow these tendencies.

2.5 Chapter Summary

In this chapter, I have discussed the grammar which defines possible NKK accent patterns vs. impossible NKK accent patterns. I suggest that the markedness constraints $H]_{PW}$, OCP(H), ALIGN-L (DOUBLE), HLINK, and WTP are undominated in NKK and play a crucial role in restricting the possible patterns to four attested patterns (initial accent, penultimate accent, final accent, and double accent on the initial two syllables). In addition, the faithfulness constraint IDENT(T) is important because lexically-given accent surfaces faithfully in the output, unless the input pattern violates one of the undominated markedness constraints. Although NKK native accent patterns are not fully predictable, certain regular tendencies emerge in loanwords: CVC syllables tend to attract accent; otherwise, penultimate accent is most common. The differences in the accent patterns of loanwords vs. native words raise the following questions:

(22) Research questions
1. What is the source of the asymmetric syllable weight distinction in loanwords?
2. How do NKK speakers accent newly adopted words, which lack accent specification in UR? Do they accent new words randomly or in certain regular patterns?
3. If some patterns emerge in the accentuation of novel words, what is the source of these patterns? Do they represent a default accent position (as has been argued for an accentless word group)?
4. If a grammar constrains accentuation of novel words, how is the grammar learnable despite incomplete evidence in the lexicon?

These research questions lead this dissertation. In order to answer these questions, experimental studies were performed using novel words to investigate what factors influence accent placement when no lexical entry is available for accent. These studies are presented in the following chapters.

Chapter 3 The Effect of Syllable Structure on Accentuation of Novel Words

3.1 Introduction

The lexical accent pattern of a word in North Kyungsang Korean (NKK) is lexically determined, as illustrated in the following minimal triple *káci* ‘kind’, *kacî* ‘eggplant’, *kácî* ‘branch’ (G. Kim 1988; Y. Chung 1991; N. Kim 1997; S. Kim 1999; Kenstowicz & Sohn 2001). Syllable structure does not predict the lexical accent position in NKK: i) CVC syllables do not necessarily attract a high tone pitch accent (e.g. *cin.tal.lé.*’azalea’; *cam.cá̋ri.*’dragon fly’; *ə́.rɨ́.sin.*’older person’); ii) all four patterns are possible in words with all light syllables (e.g. *kámani* ‘straw rice-bag’; *satarí* ‘ladder’; *apə́ci* ‘father’; *kícíke* ‘a stretch’). Given the fact that pitch accent must be lexically specified, how do NKK speakers decide to accent newly adopted words, which lack accent specification? Do they accent new words randomly or does their accent assignment reveal regular patterns? This chapter reports on an experimental study using novel words to examine accent placement when no lexical entry is available for accent.

There have been at least two different approaches to explaining patterns in novel words. One view is that patterns emerging in novel words are influenced by analogy with existing words (e.g. Albright and Hayes 2001; Eddington 2000, 2004; Face 2004). Under the analogy-based account, patterns in novel words follow familiar words which are either phonetically similar to the novel words, or which are most frequent in the lexicon. In contrast, the second approach assumes that the phonological grammar contains default rules which are applied when a lexical entry is not available. Crosswhite *et al*. (2003) showed that the dominant pattern of stress assignment by Russian speakers in novel words was stem-final stress (80%), even though the occurrence of stem-final stress is much less frequent (30%) in the lexicon. Such an approach to accentuation of Korean words has been proposed by N-J Kim (1997).

The experimental study presented in this chapter was designed to test the hypothesis that there is a default accent rule which applies when there is no lexical entry available. Among the previous studies of NKK accent patterns, N-J Kim (1997) proposes that penultimate accent is the default pattern for accentless words containing no long vowel, regardless of syllable structure. Contra the hypothesis, the results in this chapter show that while penultimate accent was the most frequent pattern in words lacking CVC syllables, CVC syllables showed a significant tendency to attract accent. Additional possible influences on accent patterns in novel words were also investigated: i) whether coda type (obstruent vs. sonorant) influences accent patterns, and ii) whether vowel quality (high vs. low) matters in accent patterns. The results showed no consistent effect of coda quality, though double accent was more frequent in the novel forms with high vowels than in the forms with low vowels. Nevertheless, the vowel quality effects did not outweigh syllable structure sensitivity.

Section 3.2 discusses the *Default Accent Hypothesis* which guided Experiment 1. Section 3.3 presents the methodology of the experimental study of NKK accentuation in novel words, and findings are summarized in Section 3.4.

3.2 Default Accent Hypothesis

Previous researchers have assumed that NKK lexical words can be divided into at least two groups: an accented word group, which has underlying specified accent, and an unaccented word group, which has no underlying accent (Y-H Chung 1991; N-J Kim 1997; S-H Kim 1999). In terms of accent patterns in the unaccented word group, N-J Kim (1997) proposes a default pattern for the placement of the high tone accent; namely, that accent falls on a word-final syllable that contains a long vowel,[10] and on the penultimate syllable otherwise. N-J Kim's penultimate default accent hypothesis is based on the following arguments: i) penultimate accent is most frequent; ii) words longer than trisyllabic get accented on the penult consistently (e.g. *kosɨmtóc^h^i* 'porcupine'); iii) penultimate accent serves as an alternative accent pattern for words whose initial syllable is accented, as shown in the following examples: *kə́puki ≈ kəpúki* 'turtle'; *kámani ≈ kamáni* 'straw rice-bag'; *kámulchi ≈ kamúlchi* 'mullet', etc.

On the other hand, S-H Kim (1999) assumes that the unaccented group gets final accent (c.f. Y-H Chung 1991). His argument is based on morphological patterning: accent shift occurs only in stems with final accent when stems occur with consonant-initial suffixes longer than monosyllabic, while stems with other accent types do not show the accent shift, as illustrated in (23).

(23) Accent shift with consonant-initial suffixes longer than monosyllabic

a. Final accented stem + suffixes

satarí 'ladder'+ coc^{h}a 'even' → satari cóc^{h}a

cintallé 'azalea' +coc^{h}a 'even'→ cintalle cóc^{h}a

b. No accent shift : non-final accented stem + suffixes

camcári 'dragon fly' + coc^{h}a 'even' → camcári coc^{h}a

kámani 'straw rice-bag'+coc^{h}a 'even'→ kámani coc^{h}a

múcíke'rainbow'+ coc^{h}a 'even' →múcíke coc^{h}a

S-H Kim argues that final accented stems are accentless in underlying representation (UR) and consonant-initial suffixes have an underlying accent. Therefore, when accented suffixes are attached to accentless stems, only the underlying accent on the suffixes surfaces. In contrast, underlying accent on the suffixes does not surface in the output when the suffixes are attached to non-final accented stems. However, his argument is undermined by patterns in loanwords. Under the assumption that loanwords are accentless in UR, S-H Kim's analysis predicts that final accented loans should show accent shift just as native final accented stems do. However, final accented loans, as well as non-final accented loans, fail to show accent shift, as shown in (24).

[10] Long vowels occur only at the edges of the stem (e.g. Y-H Chung 1991). Words in which the first syllable contains a long vowel are assigned double accent (Y-H Chung 1991: 99). However, long vowels rarely occur word-finally: i) only a few bisyllabic words are available which contain a long vowel word-finally (e.g. *apə́:m* 'father', *əmə́:m* 'mother', *halpə́:m* 'grandfather', and *maŋné:* 'the last child' are all that are found); ii) trisyllabic words (or words longer than trisyllabic) where a final syllable contains a long vowel are not attested at all.

(24) No accent shift in final accented loanwords

allatín	'Aladdin'	allatín coc^ha	'even Aladdin'
pɨracíl	'Brazil'	pɨracíl coc^ha	'even Brazil'
c.f. $ap^h\acute{a}t^h$ɨ	'apartment'	$ap^h\acute{a}t^h$ɨ coc^ha	'even apartment'

Here, the accent shift disappears in accent patterns of loanwords: accent on loan stems with either final accent or penultimate accent survives on the surface forms with suffixes. Furthermore, final accent is one of the least common patterns in native words longer than bisyllabic (e.g. 6% of trisyllabic lexical words) and is less productive in loanwords unless the words contain heavy syllables word-finally, which seems contrary to the characteristics of default accent.

3.3 Experiment 1

N-J Kim's *Default Accent Hypothesis* predicts that penultimate accent will be the most common pattern in novel words that do not contain a long vowel. In contrast, S-H Kim's hypothesis predicts that final accent will be the most common pattern in novel words. Experiment 1 was designed to test these hypotheses by asking participants to produce novel forms lacking a long vowel.

3.3.1 Participants

30 subjects, ranging in age from 29 to 54 years (mean 39), participated in this study. Participants were selected for the following characteristics: if they were born and grew up (had lived until they were 20 years old) in Daegu, a city located in the northern part of Kyungsang province; were older than 25; had completed education higher than high school level; and had no history of hearing or speaking impairment. Most of them still lived in the city at the time of the experiments. 15 were male and 15 were female. They were paid for their participation, which took less than half an hour.

3.3.2 Materials

The test word set consisted of 36 words, 20 bisyllabic and 16 trisyllabic, comprising four possible combinations: for bisyllabic words, LL, HL, LH, HH; for trisyllabic words, LLL, HLL, HHL, LLH (H:heavy; L:light) (e.g. *mana, maŋna, manaŋ, maŋnaŋ, takapa, tapkapa, tapkanpa ,takapak*). Heavy syllables used in the experiment were closed with one of the following coda consonants: *k, p, m, n, ŋ*. Light syllables were open. A single vowel, either *a* or *i*, was used for all syllables within a word (for example, *takapa; citiki*). The phonetic quality of onset consonants was also controlled: each word contained only obstruent onsets such as *p, t, k* (e.g. *kapa*) or only sonorant onsets such as *m, n* (e.g. *mana*). Sibilant consonants (*s, s', ʃ*) were avoided because the high energy of the sibilant noise might affect the pitch of the vowel following the consonant. For trisyllabic words, only obstruent consonants were used for onsets (e.g. *pakapa*, *citiki*), because sonorant onsets such as *l, m, n* trigger nasalization of obstruent coda consonants in the preceding syllable (e.g. *ta**pn**aka* → *ta**mn**aka*). Four or five items were included for each syllable weight combination. The list of all test words used in the study appears in (25).

(25) Test words

a. Bisyllabic words

CV.CV	CVC$_{s/o}$.CV	CV.CVC$_{s/o}$	CVC$_{s/o}$.CVC$_{s/o}$
mana	maŋna	manaŋ	maŋnaŋ
kapa	kakpa	kapak	kakpap
tapa	tampa	tapam	tampaŋ
cipi	cimpi	cipim	cimpin
tiki	tipki	tikik	tipkik

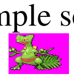
b. Trisyllabic words

CV.CV.CV	CVC$_{s/o}$.CV.CV	CV.CV.CVC$_{s/o}$	CVC$_{s/o}$.CVC$_{s}$.CV
takapa	tapkapa	takapak	tapkanpa
pakaca	paŋkaca	pakacaŋ	paŋkanca
kitici	kiktici	kiticik	kiktinci
pipici	pimpici	pipicim	pimpinci

3.3.3 Procedure

Participants were recorded in a quiet room (in the subject's house in most cases) using a cardioid condenser lavalier microphone and a Zoom H4 digital recorder at 44.1 kHz sampling rate. They read a randomized list of words given in a Korean carrier sentence from a sheet of paper, for example "*atɨri* **__________** *to satalla* k^h*atəra*" (My son asked me to buy **_________** too) with 2 repetitions. Each sheet presented 5 test words in a carrier sentence. The last word in the sheet was a filler, not used in the analysis. The second readings were analyzed. Novel words were presented with pictures of unfamiliar cartoon characters, as in (26), and introduced as the names of new cartoon characters which had recently become popular among children.

(26) Example sentence

atɨri  ***tapa*** to satalla k^hatʌra

'My son asked me to buy ***tapa*** too'

Participants were asked to read the sentences as natural conversational speech. Each speaker looked through the word list in order to familiarize himself/herself with all the words before recording. Speakers were asked to read in their own dialectal accent.

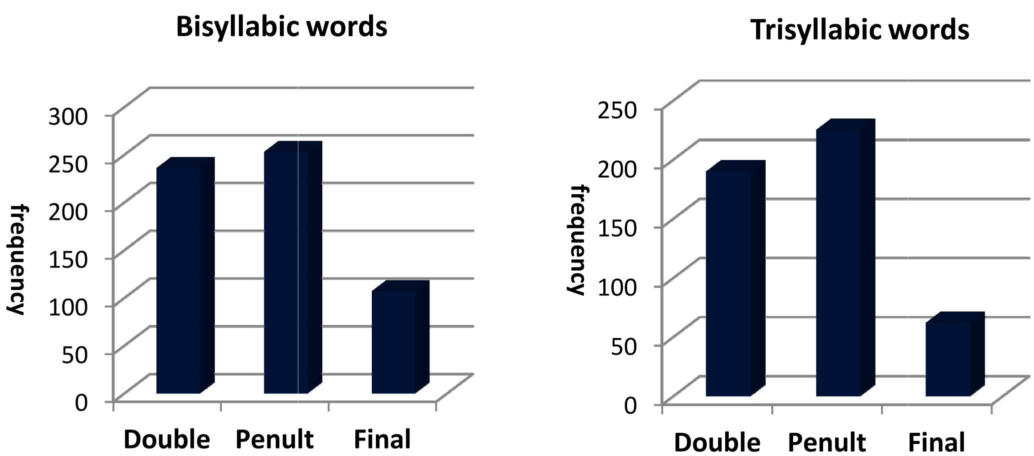
Accent patterns in target words were judged by two native NKK speakers (the author and an undergraduate linguistics student at Stony Brook University). If disagreement occurred between the two judgments, the accent decision was made based on the pitch contour generated using Praat.

3.3.4 Results

3.3.4.1 The effect of syllable structure

A total of 1073 tokens were collected from 36 test words produced by 30 subjects. 7 tokens were excluded because they contained production errors,[11] leaving 596 bisyllabic and 477 trisyllabic tokens. The *Default Accent Hypothesis* predicts that either penultimate accent or final accent should be the most frequent, regardless of syllable structure. As shown in Figure 1, penultimate accent was the most frequent in bisyllabic and trisyllabic words: 42% (253/596) in bisyllabic words; 47% (225/477) in trisyllabic words. However, double accent was assigned frequently as well: 40% of bisyllabic words and trisyllabic words (236/596; 190/477) were given double accent. Final accent was least favored (less than 20%) in both groups (107/596(18%) in bisyllabic words; 62/477(13%) in trisyllabic).

Figure 1 Overall results in bisyllabic words and in trisyllabic words



Bisyllabic words

Table 1 presents counts of each accent according to syllable structure combinations in bisyllabic words.

Table 1 Distribution of accent for bisyllabic words according to syllable structure combinations

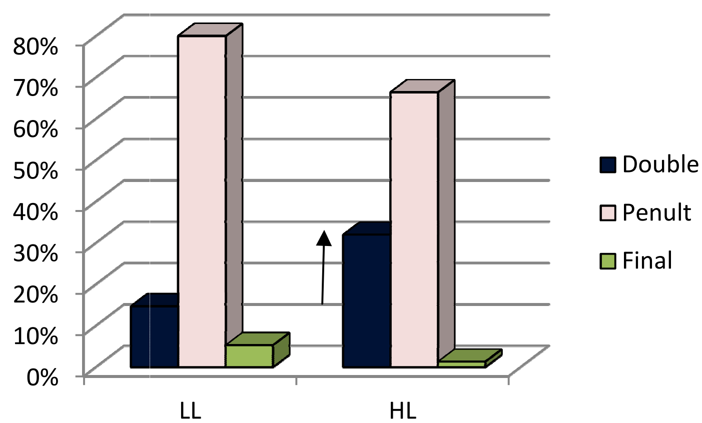
Word Type	Accent Type			Total
	Double (%)	Penult (%)	Final (%)	
LL	22 (14.7)	**120 (80)**	8 (5.3)	150
HL	48 (32)	**99 (66.4)**	2 (1.3)	149
LH	42 (28)	23 (15)	**84 (56)**	149
HH	**124 (84)**	11 (7.4)	13 (8.8)	148
Total	236 (40)	**253 (42)**	107 (18)	596

[11] The example tokens excluded due to production errors were as follows: *kakpap* produced for the target word *kakpa* (S2); *manaŋ* for *maŋnaŋ* (S2); *tampa* for *tapam* (S7); *kintinci* for *kikdinci* (S7); *taŋkanpa* for *tapkapa* (S11).

The bolded numbers indicate the most frequent accent type for each combination in Table 1. Penultimate accent was the most common accent pattern in LL and HL words. However, final accent was most frequent in LH words and double accent was most frequent in HH words. Therefore, the distribution of accent according to each syllable weight combination showed that accent assignment varies according to syllable weight: penultimate accent was predominant in LL but rare in LH and in HH; final accent was least favored across the combinations except in words with a word-final heavy syllable (LH); and double accent was overwhelmingly common in words where the initial two syllables were heavy (HH). The weight of the final syllable seemed to be crucial in inducing non-penultimate accent.

Statistical analyses using Pearson's chi-square confirmed that the association between accent type and syllable structure combination was statistically significant. The independent variables used were syllable structure combination and accent position. The dependent variable was the number of tokens. Comparing LL and HL shows that penultimate accent was chosen 80% of the time in words with all light syllables (LL) and 66% in words with initial heavy syllables (HL). However, as indicated by the arrow[12] in Figure 2, double accent placement was more common (32%) in HL than in LL, resulting in significant differences in the occurrence of double accent in HL vs. LL, with 15% double accent ($\chi^2(2)=15.9$, $p<.001$). The accent patterns in HL vs. LL are illustrated in Figure 2.

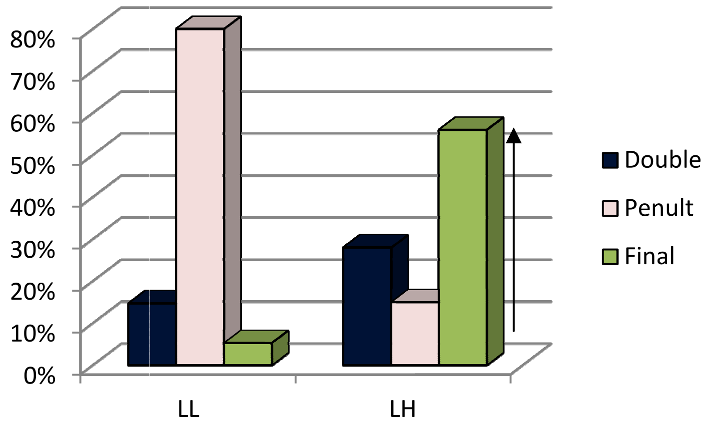
Figure 2 Accent distribution in LL vs. in HL



Comparing LH accent distribution with LL, as shown in Figure 3, final accent was assigned in 53% of the words with a final heavy syllable (LH). This was significantly different from patterns of LL, which had final accent in only a few words (5 %) ($\chi^2(2)=133$, $p<.001$). In contrast, 15% of LH forms received penultimate accent, significantly less than the 80% of penultimate accent in LL.

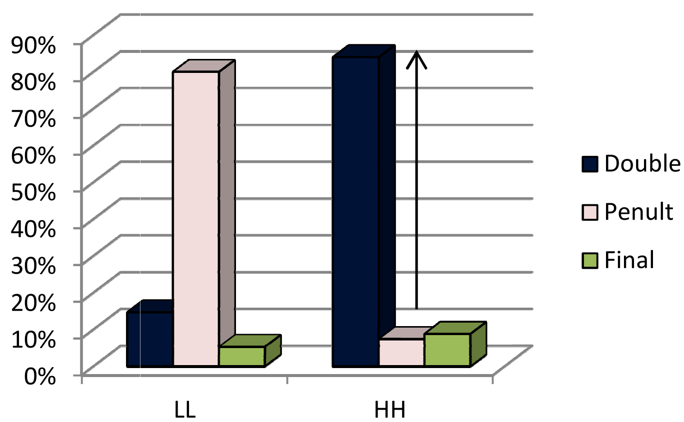
[12] The up arrow in the figures throughout the chapter indicates the notable difference in accent distribution from the other group.

Figure 3 Accent distribution in LL vs. in LH



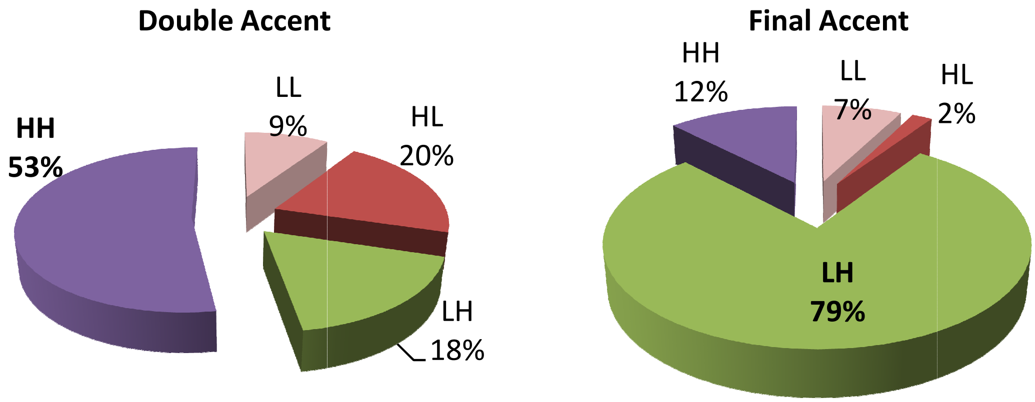
In terms of accent distribution of HH combinations, it is striking that double accent was overwhelmingly prevalent in HH. 124 out of 148 words (84%) were given double accent while penultimate accent and final accent were given in only a few words, 7% and 9%, respectively. As described in Figure 4, the dominance of double accent in HH contrasts with the dominant penultimate accentuation in LL. Pearson's Chi-square analyses also confirmed that the accent distribution of HH was statistically different from that of LL ($\chi^2(2)=165$, $p<.001$).

Figure 4 Accent distribution in LL vs. in HH



In Figure 5, the distribution of double accent and final accent according to syllable structure combinations demonstrates structure-sensitive accentuation more clearly.

Figure 5 Structure-sensitive accentuation in bisyllabic words



Among the words with double accent, the HH combination was the most favored and the percentage (53%) was higher than that of other combinations (LL, HL, LH). In addition, most of the final accented words were LH. The probability of final accent in LH (79%) was much higher than other combinations. Even though final accent was not preferred in general, final accent was favored most when the word-final syllable was heavy.

Trisyllabic words

For trisyllabic words, the results were similar to those of bisyllabic words: penultimate accent was most frequent, double accent less, and final accent was least favored. However, double accent was also common overall. The results are shown in Table 2.

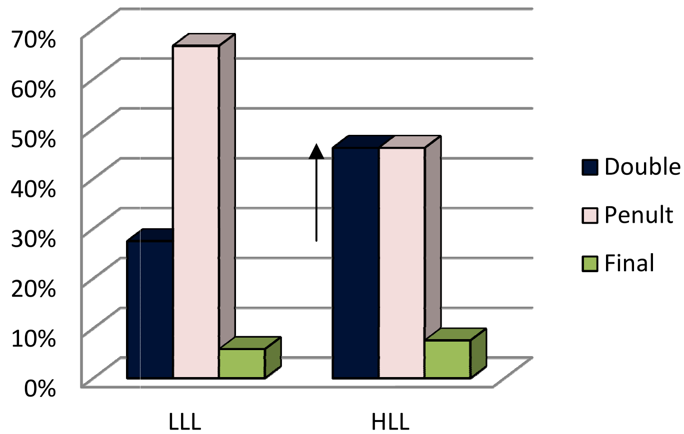
Table 2 Distribution of accent in trisyllabic words according to syllable weight combinations

Word Type	Accent Type			Total
	Double (%)	Penult (%)	Final (%)	
LLL	33 (28)	**80 (67)**	7 (6)	120
HLL	**55 (46)**	**55 (46)**	9 (8)	119
LLH	35 (29)	**45 (38)**	**40 (33)**	120
HHL	**67 (57)**	45 (38)	6 (5)	118
Total	190 (40)	**225 (47)**	62 (13)	477

As indicated by bolded numbers in Table 2, the most frequent accent type varies according to syllable weight. Penultimate accent was most frequent in LLL, HLL and LLH, though less common in HLL and LLH than in LLL. In contrast, double accent was more common (46%) in HLL and final accent (33%) in LLH. Double accent was most frequent in HHL (57%).

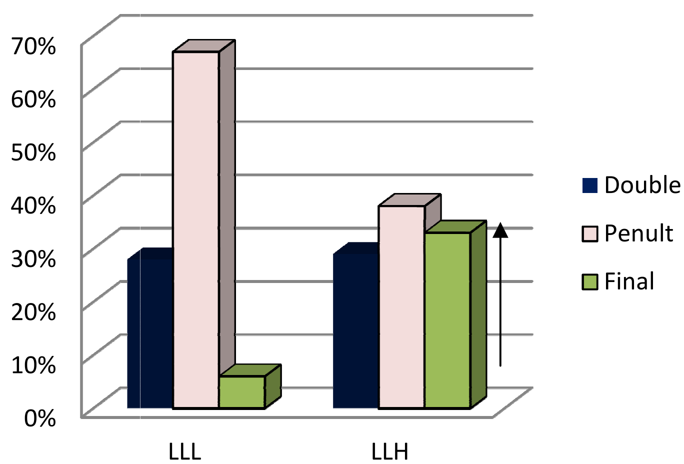
The distribution of HLL, LLH and HHL was compared with that of LLL using Pearson's Chi-square analyses. As shown in Figure 6, in comparison of HLL and LLL, HLL has less frequent penultimate accent (46%) and more frequent double accent (46%) than those in LLL (67%, 28% respectively), and a difference in distribution was statistically significant ($\chi^2(2)=10.7$, $p=.005$).

Figure 6 Accent distribution in LLL vs. in HLL



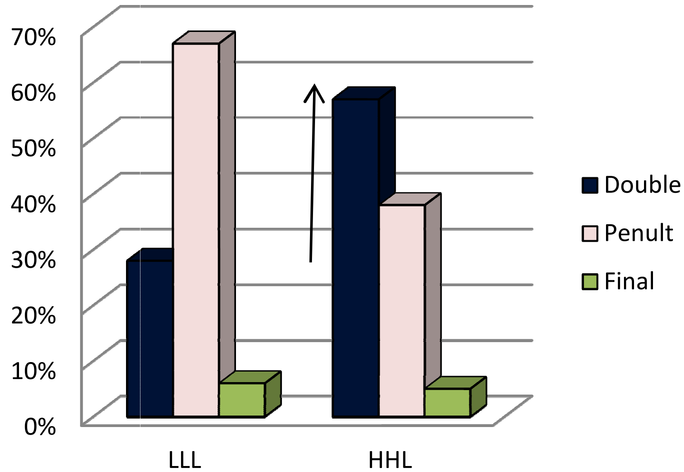
In terms of accent patterns of LLH, three accent types were relatively evenly distributed, as shown in Figure 7: double accent 29%; penultimate accent 38%; final accent 33%. The tendency toward final accent for LLH (33%), however, was significantly stronger than the tendency toward final accent in other combinations, since final accent was fairly rare in other combinations (less than 10%). Statistical analyses confirmed that the accent distribution of LLH was significantly different from that of LLL ($\chi^2(2)=33$, $p<.001$)

Figure 7 Accent distribution in LLL vs. in LLH



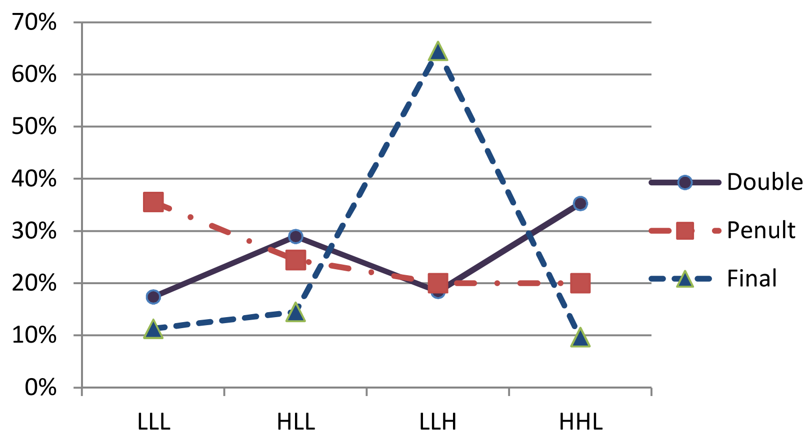
In Figure 8, double accent was the most frequent pattern (57%) in HHL. Penultimate accent was less common in HHL (38%), a significantly different distribution from that of LLL ($\chi^2(2)=22.6$, $p<.001$).

Figure 8 Accent distribution in LLL vs. in HHL



The structure-sensitive accentuation was revealed more clearly in the distribution of each accent according to syllable weight combinations, as illustrated in Figure 9.

Figure 9 Structure-sensitive accentuation in trisyllabic words



Double accent was more frequent where initial syllables were heavy (HLL and HHL) and final accent was preferred in LLH even though it was rare in all other combinations. Penultimate accent was less preferable in HLL, LLH, and HHL.

To summarize, penultimate accent was the most frequent both in bisyllabic and in trisyllabic novel forms, as predicted by N-J Kim's claim that penultimate accent represents the default pattern. However, penultimate accent was less favored in words which contained heavy syllables word-initially or word-finally. Therefore, the syllable structure-sensitive tendencies in novel words do not support the hypothesis that speakers confronted with novel forms will assign accent in accord with a default preference for accent in penultimate position, regardless of syllable structure.

3.3.4.2 Effects of coda type

The previous section considered whether syllable type, open vs. closed, influenced the position of accent in novel words. This section examines whether the quality of coda consonants, e.g. sonorant vs. obstruent, influences accent patterns in novel words. Kenstowicz and Sohn (2000) proposed the syllable weight scale in (27) to account for NKK accent placement in loanwords. According to the syllable weight scale, long vowels are the heaviest; syllables with sonorant coda consonants are heavier than syllables with obstruent coda consonants; and lexical vowels are heavier than epenthetic vowels. In terms of coda type effects on accent patterns in novel words, the syllable weight scale predicts that heavy syllables with sonorant codas should attract accent more commonly than syllables with obstruent codas.

(27) $V\!: > VC_{son} > VC_{obs} > V > V_{epen}$

The accent distribution in bisyllabic HL, LH, and HH novel forms according to coda quality is presented in Table 3.

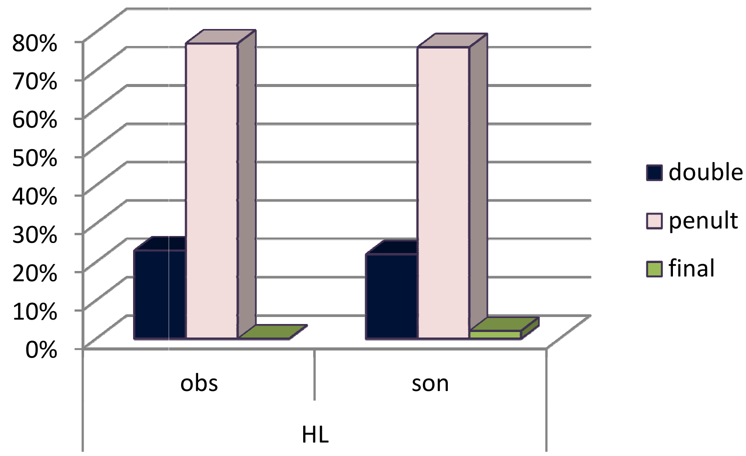
Table 3 Accent distribution in bisyllabic words: obstruent codas (obs) vs. sonorant codas (son)

Word Type	Coda type in H	Double(%)	Penult(%)	Final(%)	Total (%)
HL	obs	29 (48)	**31(52)**	0(0)	60(100)
	son	20(22)	**68(76)**	2(2)	90(100)
LH	obs	15(25)	18(30)	**27(45)**	60(100)
	son	27(30)	6(7)	**57(63)**	90(100)
HH	obs	**54(90)**	4(7)	2(3)	60(100)
	son	**72(80)**	7(8)	11(12)	90(100)

In general, obstruent codas in word initial syllables attracted double accent more frequently than sonorant codas (48% vs. 22% in HL; 90% vs. 80% in HH). On the other hand, sonorant codas in word final syllables attracted final accent more often than obstruent codas (63% vs. 45% in LH; 12% vs. 3% in HH).

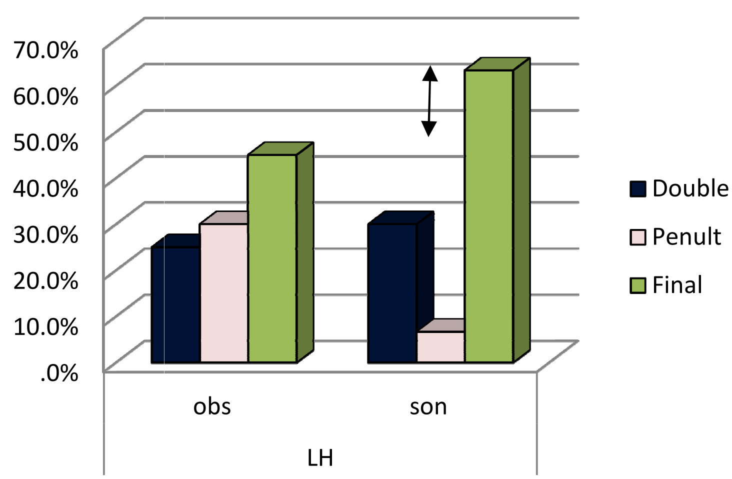
As shown in Table 3, penultimate accent was the most frequent in HL both with obstruent codas and with sonorant codas (52%; 76%). However, penultimate accent was less frequent when the initial syllable was closed by obstruent codas. Instead, double accent was assigned more frequently in HL with obstruent codas than with sonorant codas (48% vs. 22%). However, item analyses revealed that the different distribution was not related to the coda quality, because the difference actually resulted from the peculiar accent patterns of a specific item: *kakpa* was given double accent in 72% of responses (21/29), while the other HL items were given double accent on average 23%: *tampa* 7/30; *cimpi* 5/30; *tipki* 7/30; *maŋna* 8/30. After the item *kakpa* was excluded from the statistics, the accent distribution in HL with obstruent codas vs. with sonorant codas was not much different, as illustrated in Figure 10. (The item analyses are discussed further in section 3.3.4.3.)

Figure 10 Obstruent coda vs. Sonorant coda in HL



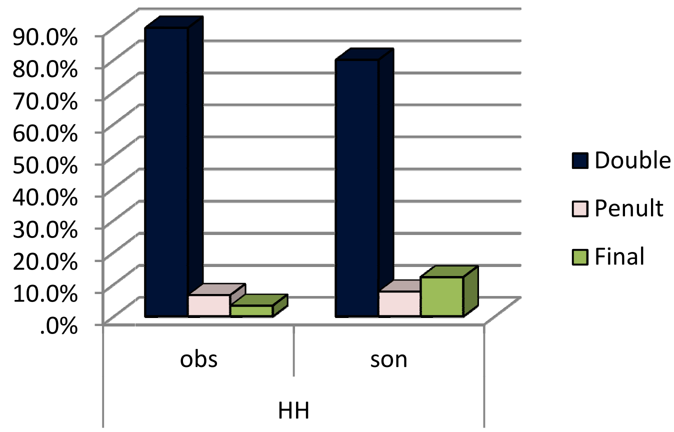
On the other hand, final accent was most common in LH both with obstruent codas and with sonorant codas, as shown in Figure 11. As predicted by the syllable weight scale (27), word-final syllables with sonorant codas are more likely to attract accent than syllables with obstruent codas. The difference in accent distribution in LH forms with obstruent codas vs. with sonorant codas was statistically significant ($\chi^2(2)=14.7$, p=.001).

Figure 11 Obstruent coda vs. Sonorant coda in LH



However, the accent patterns of HH were not much different for obstruent codas vs. sonorant codas, as illustrated in Figure 12. Double accent was predominant in both groups of HH. The assignment of double accent was more frequent in HH with obstruent codas than with sonorant codas. In contrast, final accent was more frequent in HH with sonorant codas than in HH with obstruent codas. However, the difference was not statistically significant ($\chi^2(2)=3.8$, p=.152).

Figure 12 Obstruent coda vs. Sonorant coda in HH



For trisyllabic words, coda type effects were not significant across all the combinations (p>.05): the accent distribution in HLL/HHL was not much different according to the coda quality, whereas final accent in LLH with sonorant codas was somewhat more frequent than with obstruent codas. The results are given in Table 4.

Table 4 Accent distribution in trisyllabic words: obstruent codas vs. sonorant codas

	Coda type	Double(%)	Penult(%)	Final(%)	Total(%)
HLL	obs	**27(45)**	29(48)	4(7)	60(100)
	son	**28(47)**	26(43)	6(10)	60(100)
LLH	obs	18(30)	24(40)	**18(30)**	60(100)
	son	17(28)	21(35)	**22(37)**	60(100)
HHL	obs	**34(57)**	21(35)	5(8)	60(100)
	son	**35(58)**	24(40)	1(2)	60(100)

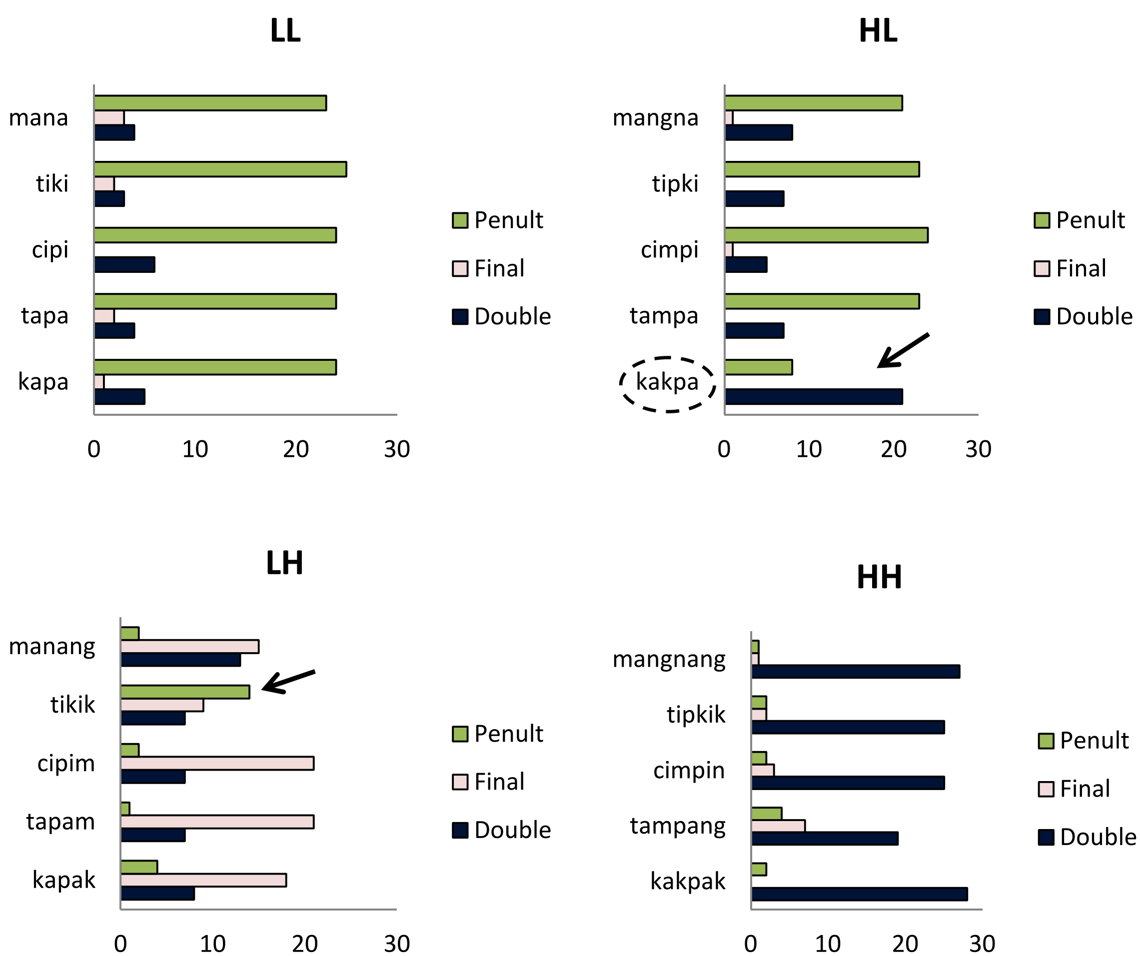
To summarize, heavy syllables with sonorant codas attracted accent more often than heavy syllables with obstruent codas in bisyllabic LH words, in accordance with the predictions of the syllable weight scale (27). However, accent patterns in other word types in bisyllabic forms and in trisyllabic forms did not show a consistent effect of coda quality. Therefore, I conclude that the syllable weight scale is not strictly observed when NKK speakers accent novel words.

3.3.4.3 Item effects

Item analyses are necessary to confirm that the accent patterns reported so far indeed represent the typical accent assignment for each word type.

First, specific patterns of each item in bisyllabic words were presented in Figure 13. For each syllable structure combination, five test items were used.

Figure 13 Item analyses in bisyllabic novel forms: accent distribution in each item

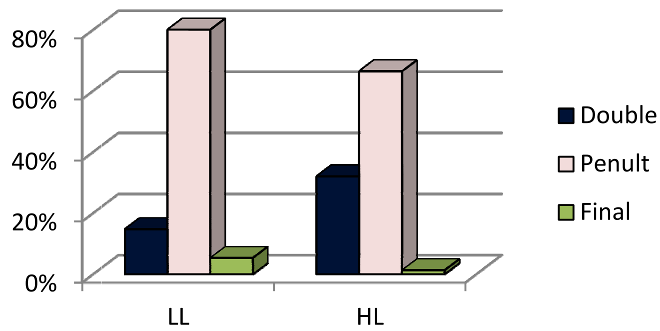


As shown in Figure 13, accent patterns in all items in LL combinations were consistent, showing predominantly penultimate accentuation. However, in HL combinations, *kakpa* was peculiar in that double accent was predominant (double 72%), whereas penultimate accent was more common in the remaining HL items (double 23% vs. penult 76%).[13] Recall that the accent distribution in HL novel forms was significantly different from that in LL forms. This distinct distribution was ascribed to more frequent double accent and less frequent penultimate accent in HL than in LL, as illustrated in Figure 2. The figure is repeated below:

---

[13] As for a question of why *kakpa* should have been treated differently, I found that most existing bisyllabic *sino*-Korean words beginning with *kak* tend to carry double accent (e.g. *káktó* 'angle'; *káksə́* 'memorandum'; *kákcá* 'each one' c.f. *pure*-Korean *káksi* 'a bride'). I speculate that the different accent pattern in *kakpa* might be related to the association of double accent with the syllable *kak-* in the existing words.

Figure 14 Accent distribution in LL vs. in HL



However, the statistically significant effects in accent assignment in HL vs. in LL actually resulted from a single item, *kakpa*. Excluding the item *kakpa* from the statistical analyses, the accent patterns in HL vs. LL were not much different, as shown in Table 5. Penultimate accent was the most common pattern in both word types (76% in HL; 80% in LL), whereas double accent was somewhat more frequent in HL than in LL (23% in HL vs. 15% in LL). However, the difference in accent patterns of HL from LL did not reach statistical significance ($\chi^2(2)=4.8$, $p=.09$). The accent distribution in LL vs. in HL excluding *kakpa* is presented as follows:

Table 5 Accent distribution in LL vs. in HL (kakpa excluded)

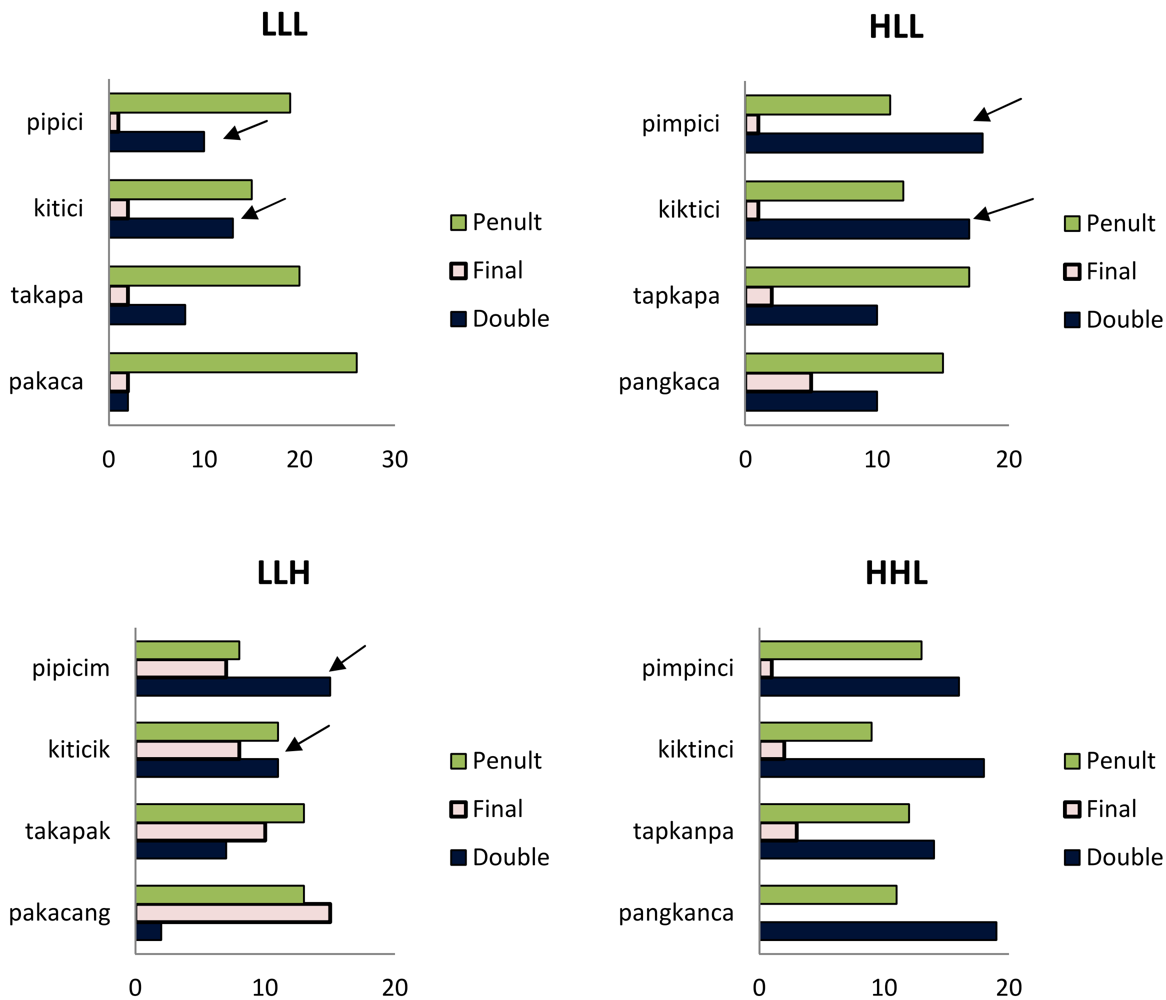
	Double	Penult	Final	Total
LL	22/150(14.7%)	**120/150(80%)**	8/150(5.3%)	150
HL	27/120(22.5%)	**91/120(75.8%)**	2/120(1.7%)	120

This result suggests that word-initial heavy syllables in bisyllabic novel words do not increase penultimate (HL) accentuation due to the competing effect of double accent for HL. Nevertheless, the preference for double accent in HL was not strong enough to bring about significant difference in the overall tendencies.

In LH combinations, all the words except *tikik* showed strong final accent: 63% final accent for all forms excluding *tikik* vs. 56% including *tikik*. Finally, in HH, accent patterns of all the individual items were consistent: double accent was the most common pattern in HH.

Trisyllabic novel forms also revealed item specific variations in accent patterns, as illustrated in Figure 15.

Figure 15 Item analyses in trisyllabic novel forms: accent distribution in each item



Despite dominant penultimate accent in LLL, double accent was more common in items containing a high vowel *i* (*pipici*; *kitici*) than those with a low vowel *a* (*takapa*; *pakapa*). The tendency toward double accent in words with a high vowel persisted in HLL combinations: items with a high vowel, *pimpici* and *kiktici*, were given double accent more frequently than items with a low vowel, *tapkapa* and *pangkaca* (58% vs. 34%). In LLH combinations, double accent was also preferred in the items with high vowels (43% in the items with high vowels vs. 15% in the items with low vowels): double accent in *pipicim* and *kiticik* was more common, while double accent was fairly weak in the other words. Penultimate accent and final accent were stronger in the other items (*takapak*; *pakacang*). Nevertheless, these item effects did not outweigh the syllable structure effects in novel forms. The effects of syllable structure were also present in accent patterns in the items with high vowels (e.g. *pipici*; *pimpici*; *pipicim*; *pimpinci*) ($\chi^2(6) = 29.08$, $p < .001$, $N = 240$). Finally, as for HHL combinations, no specific item was found to show distinct patterns from the general tendency toward accented heavy syllables.

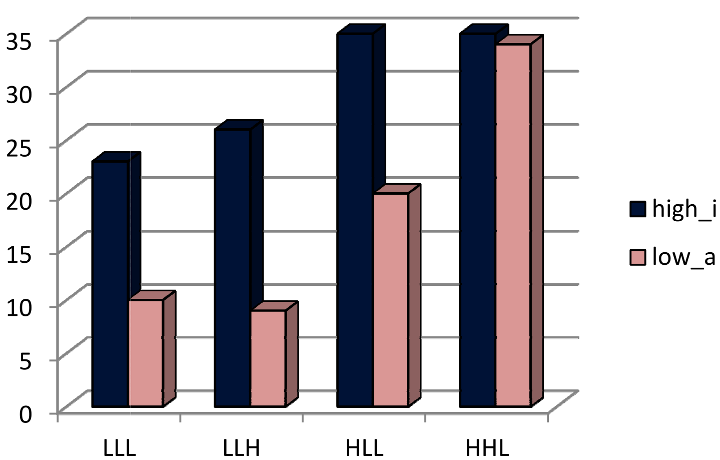
Item specific patterns in trisyllabic forms revealed that novel forms containing a high vowel *i* were more likely to have double accent. The preference for double accent in words with a high vowel persisted even in cases where the double accent contradicted the syllable structure-sensitive tendencies; e.g., double accent was common in *pipicim* and *kiticik*. The vowel quality effects will be examined more systematically in the following section, accompanying statistical analyses.

3.3.4.4 Vowel quality effects

This section examines vowel quality effects on the accent patterns in novel words. In each syllable structure combination, accent patterns in novel forms containing a high vowel *i* were compared to those in novel forms containing a low vowel *a*. Since the previous section has already shown that double accent was more likely in items with high vowels, only the assignment of double accent was compared for the vowel quality effects. In addition, Pearson's chi-square analyses were performed to determine whether the vowel quality effects on the accent patterns were significant. The independent variable was vowel quality (high; low) and the dependent variable was the number of tokens with double accent.

As shown in Figure 16, the frequency of tokens with double accent was compared for each syllable structure combination. The two bars in each word type indicate the frequency of double accent: in forms with high vowels and in forms with low vowels.

Figure 16 Double accent in novel forms: with high vowels vs. with low vowels



Overall, double accent was more likely in HLL and HHL than in LLL and LLH, both with high vowels and with low vowels. Furthermore, double accent was consistently more frequent in novel forms containing high vowels than in novel forms containing low vowels: 38% vs.17% in LLL; 43% vs. 15% in LLH; 58% vs. 33% in HLL; 58% vs. 57% in HHL. The statistical tests confirmed that the vowel quality effects were significant in all the combinations except HHL: in LLL, $\chi^2(2)=7.1$, $p=.03$; in LLH, $\chi^2(2)=11.8$, $p=.003$; in HLL, $\chi^2(2)=9.2$, $p=.01$. The accent patterns in HHL, however, were not much different according to the vowel quality ($\chi^2(2)=.04$, $p>.05$).

In sum, vowel quality indeed influenced accent patterns in novel words, which might serve as a factor in accent variation in trisyllabic novel words. Double accent was more frequent in the novel forms with high vowels than in the forms with low vowels. Nonetheless, the effects did not override syllable structure sensitivity. I conjecture that the preference for double accent in the forms with high vowels could be ascribed to the inherent phonetic quality (high pitch) of the high vowels: high vowels such as [i] and [u] tend to have higher F0 than low vowels such as [a] (e.g. C-W Kim 1968; Ladd & Silverman 1984; Whalen and Levitt 1995; Yang 1996; Hoole *et al*. 2011). However, only the word-initial high vowels are important because only the word-initial high vowel (but not the word-final high vowel) was more likely to be associated with a high tone. Double accent rather than final accent was preferred in the forms where the word-final vowel as well as the word-initial vowel was high; when the word-initial vowel was high, double accent was common even in cases where final syllables were heavy (e.g. LLH forms with high vowels).

3. 4 Chapter Summary

This chapter has presented an experimental study of NKK accent patterns in novel forms, which tested the *Default Accent Hypothesis*. The results showed that NKK speakers tended to accent novel forms not randomly, but in a patterned manner: penultimate accent was the most common pattern, both in bisyllabic and in trisyllabic novel forms; double accent was more likely in words beginning with heavy syllables (HLL; HHL); final accent was more common in words with final heavy syllables (LH; LLH). These patterned tendencies are consistent with N-J Kim's default penultimate accent hypothesis, which predicted penultimate accent as default for unaccented words without a long vowel, but not with S-H Kim's final accent hypothesis for unaccented words. However, the results also showed that not only penultimate accent but also other NKK accent patterns (double accent; final accent) were possible for unaccented novel words. Availability of double accent in novel words actually contradicts the predictions of the two hypotheses. Therefore, a preference for a default accent position, either penultimate or final, fails to predict the syllable structure effects and variation in accent patterns in novel words. In addition, the absence of the initial accent pattern in novel words suggests that the initial accent pattern is indeed marked and not grammatically motivated.

In terms of the NKK lexical accent patterns, CVC syllables do not always attract a pitch accent (e.g. *cin.tal.lé.*'azalea'; *cam.cá.ri.*'dragon fly'). On the contrary, syllables with long vowels always attract accent (e.g. *maŋné:* 'the youngest'; *hó:ráŋi* 'tiger'). Nevertheless, existence of coda consonants clearly has an effect on accent placement of novel words: word-initial/word-final CVC syllables *tend* to attract accent. Chapter 6 will discuss possible analyses of the difference between native words and novel words.

In addition, accent patterns in novel forms with sonorant codas vs. obstruent codas were compared in order to examine coda type effects on accent patterns in novel words. In accordance with the syllable weight scale proposed by Kenstowicz and Sohn (2000), CVC syllables with sonorant codas should attract accent more commonly than CVC syllables with obstruent codas. However, against the prediction, the results showed that coda quality did not consistently influence the accent placement in novel words: CVC syllables with sonorant codas attracted accent more frequently than CVC syllables with obstruent codas in bisyllabic LH novel forms but not either in other bisyllabic novel forms or in trisyllabic novel forms. On the other hand, vowel quality did influence accent assignment in novel words: word-initial high vowels were more likely to attract accent than word-initial low vowels. I assume that this effect might be ascribed to the intrinsic high pitch of high vowels (e.g. C-W Kim 1968; Ladd & Silverman 1984; Whalen and Levitt 1995): a pitch accent is more likely to be associated with elements with high pitch. This suggests that phonetic qualities would serve as one of the factors associated with accent variations.

To summarize, we found the following generalizations in NKK accent patterns that emerged in novel words:

(28) NKK accent patterns in novel words

- penultimate accent was the most common pattern, both in bisyllabic and in trisyllabic novel forms
- double accent was more likely in words beginning with heavy syllables (HLL; HHL)
- final accent was more common in words with final heavy syllables (LH; LLH)
- CVC syllables with sonorant codas did not consistently attract accent more often than CVC syllables with obstruent codas
- word-initial high vowels were more likely to attract accent than word-initial low vowels

Finally, I argue that the structure-sensitive accent patterns emerging in novel words are grammatically motivated. This raises the question of where the grammatically motivated patterns originate. As discussed earlier, syllable structure does not determine the accent position categorically in NKK native accentuation. In order to answer this question, I investigate the lexical statistics of accent patterning in NKK in the following chapter.

Appendix I: Results of Experiment 1

Table 6 Results in bisyllabic words in Experiment 1

Syll.Strctr	Novel word	Accent in novel words			Total
		Double	Final	Penult	
LL	kapa	5	1	24	30
	tapa	4	2	24	30
	cipi	6	0	24	30
	tiki	3	2	25	30
	mana	4	3	23	30
HL	kakpa	21	0	8	29
	tampa	7	0	23	30
	cimpi	5	1	24	30
	tipki	7	0	23	30
	mangna	8	1	21	30
LH	kapak	8	18	4	30
	tapam	7	21	1	29
	cipim	7	21	2	30
	tikik	7	9	14	30
	manang	13	15	2	30

HH	kakpak	28	0	2	30
	tampang	19	7	4	30
	cimpin	25	3	2	30
	tipkik	25	2	2	29
	mangnang	27	1	1	29
Total		236	107	253	596

Table 7 Results in trisyllabic words in Experiment 1

Syll.Strctr	**Novel word**	**Accent in novel words**			**Total**
		Double	**Final**	**Penult**	
LLL	pakaca	2	2	26	30
	takapa	8	2	20	30
	kitici	13	2	15	30
	pipici	10	1	19	30
HLL	pangkaca	10	5	15	30
	tapkapa	10	2	17	29
	kiktici	17	1	12	30
	pimpici	18	1	11	30
LLH	pakacang	2	15	13	30
	takapak	7	10	13	30
	kiticik	11	8	11	30
	pipicim	15	7	8	30
HHL	pangkanca	19	0	11	30
	tapkanpa	14	3	12	29
	kiktinci	18	2	9	29
	pimpinci	16	1	13	30
Total		190	62	225	477

Chapter 4 The Effect of Lexical Frequency

4.1 Introduction

Chapter 3 presented evidence that syllable structure plays a role when NKK speakers accent novel words: word-initial/final syllables were likely to be accented when they were heavy, while penultimate syllables were accented most frequently when the word did not have a heavy initial or final syllable. The syllable structure-sensitive tendencies emerging in novel words raise the question of how NKK speakers have acquired these structure-sensitive patterns, given the relatively large number of forms in the native lexicon that do not conform to these generalizations.

This chapter presents a corpus study of lexical statistical NKK accent patterns, designed in order to investigate possible sources of the syllable structure-sensitive accent tendencies emerging in novel words.

In this chapter, I propose the *Stochastic Accent Hypothesis* which claims that NKK speakers' accentuation of novel words is based on a stochastic grammar which reflects the patterns in the existing NKK lexicon. Many previous works have shown that statistical patterns in the lexicon match well with speakers' intuition (e.g. Coleman and Pierrehumbert 1997; Zuraw 2000; Frisch and Zawaydeh 2001; Ernestus and Baayen 2003; Hayes and Londe 2006). Zuraw (2000), for example, presented evidence from an acceptability-judgment experiment that Tagalog native speakers' judgments of novel words were consistent with lexical frequency of existing words with the same pattern. She argues that speakers of Tagalog construct a stochastic grammar in which lower-ranked markedness constraints are often violated by lexicalized words. These lower-ranked markedness constraints, the "subterranean" grammar, are active in the production of newly adopted words. Hayes and Londe (2006) provided additional evidence for sensitivity to lexical frequency from a study of Hungarian vowel harmony. They showed that the vowel harmony patterns in novel words were consistent with frequency in the Hungarian lexicon. This implies that Hungarian native speakers know the frequency of competing harmony patterns and they use this knowledge when they produce novel words.

The *Stochastic Accent Hypothesis* predicts that the syllable structure-sensitivity exhibited by our NKK participants emerges from the statistical patterning in the lexicon. If this is the case, we expect that while the accent patterns of existing lexical items are not fully predictable, study of the lexicon will reveal statistical tendencies that are in accordance with the NKK accent patterns in novel words. A corpus study of lexical statistics was performed to test this hypothesis. This chapter will show that syllable structure is indeed important in NKK lexical statistical accent patterning as in the structure-sensitive accent tendencies in novel words. This result suggests that NKK speakers construct a stochastic grammar based on lexical frequency. Furthermore, learning this grammar is also possible via the native lexicon and this grammar becomes active when accenting newly adopted words.

Section 4.2 presents the lexical distribution of syllable structure combinations. Section 4.3 presents the lexical statistics of the accent distribution according to syllable structure. Section 4.4 discusses how the accent patterns in novel words coincide with lexical statistical patterning. Section 4.5 concludes.

4.2 Lexical Distribution according to Syllable Structure

The *Stochastic Accent Hypothesis* predicts that the syllable structure-sensitive patterns emerging in novel words are actually consistent with statistical patterning in the lexicon. In order to examine lexical statistics, I constructed a corpus based on the frequency data of current Korean word usage (The National Institute of the Korean Language 2003; www.korean.go.kr), which consists of lexical words used in daily life and reports on frequency of use. I assume that this database is representative of the practical lexicon commonly accessed by Korean speakers. The corpus was composed of accent-indicated written forms, of which lexical items were from the database. Since the purpose of this study was to look for statistical patterns in the pure native lexicon, Sino- Korean words and loanwords were excluded from the data. In order to control morphological effects on accentuation, only monomorphemic words were selected, excluding compound nouns and derived nouns. Finally, 1275 bisyllabic and trisyllabic monomorphemic native Korean nouns that remained after excluding loanwords, Sino-Korean words, derived nouns, words longer than trisyllabic, etc were chosen out of 39,856 nouns. Among 1275 words, 865 were bisyllabic and 410 were trisyllabic words.

First of all, lexical distribution was examined according to syllable structure combinations. The word counts for each syllable structure combination are as follows:

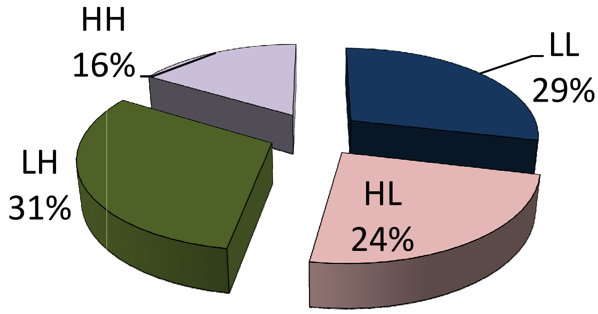
Table 8 Distribution of lexical words according to syllable structure combinations

	Bisyllabic words (865)				Trisyllabic words (410)							
Word Type	LL	HL	LH	HH	LLL	HLL	LHL	LLH	HHL	LHH	HLH	HHH
Freq.	248	205	268	144	164	88	56	25	60	9	3	5
%	28.7	23.7	31.0	16.6	40.0	21.5	13.7	6.1	14.6	2.2	0.7	1.2

*H: heavy syllable(CVC); L: light syllable(CV)

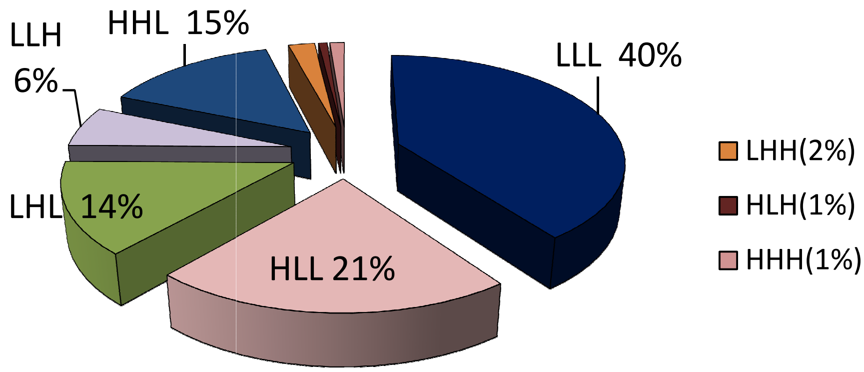
In bisyllabic words, the different syllable structure combinations are relatively evenly distributed, as shown in Figure 17: about 29% for words with all open syllables (LL); 24% for words with initial closed syllable and final open syllable (HL); 31% for words with initial open syllable and final closed syllable (LH); 17% for words with all closed syllables (HH).

Figure 17 Lexical distribution of bisyllabic words according to syllable structure combinations



However, in trisyllabic words, the distribution of syllable structure combinations is somewhat skewed: 40% of trisyllabic words contain only open syllables (LLL); 22% contain an initial closed syllable (HLL); 15% contain closed syllables in first and second position (HHL), 14% a closed second syllable (LHL), and the remaining 10% consist of other syllable structure combinations (LLH, LHH, HLH, HHH). Thus, LLL, HLL, HHL, and LHL are the major types, with LLL combination the most frequent. HLL is also relatively common compared to other combinations.

Figure 18 Lexical distribution of trisyllabic words according to syllable structure combinations



4.3 Accent Distribution according to Syllable Structure

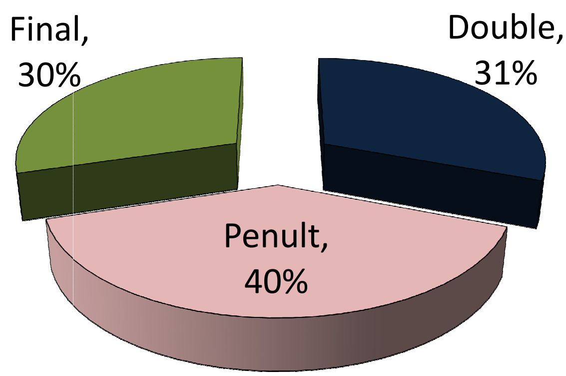
This section presents the lexical frequency of accent patterns according to syllable structure combinations. Since accent is not represented in orthography and the database does not indicate accent, the accent patterns were determined based on the intuitions of three native NKK speakers. The accent patterns given by these speakers agreed in 92% of the corpus (1172/1275): 93% (807/865) of the bisyllabic words and in 89% (365/410) of the trisyllabic words. As for words with variant accent patterns, the pattern assigned by 2 out of 3 speakers was selected.[14]

4.3.1 Bisyllabic words

As shown in Figure 19, the bisyllabic words show a relatively balanced accent distribution: 30% of the words have double accent, about 40% penultimate accent, and 30% final accent.

[14] Variant accent patterns were given in some words (8% of the corpus): 58 out of 865 of the bisyllabic words and 45 out of 410 trisyllabic words (e.g. hwíp^{h}áram ~ hwip^{h}arám 'whistle'; tórjə́nnim ~ torjə́nnim 'brother-in-law').

Figure 19 Accent distribution in bisyllabic words



However, the accent patterns are not consistent across the syllable structure combinations. The distribution of accent according to syllable structure is presented in Table 9 and Figure 20.

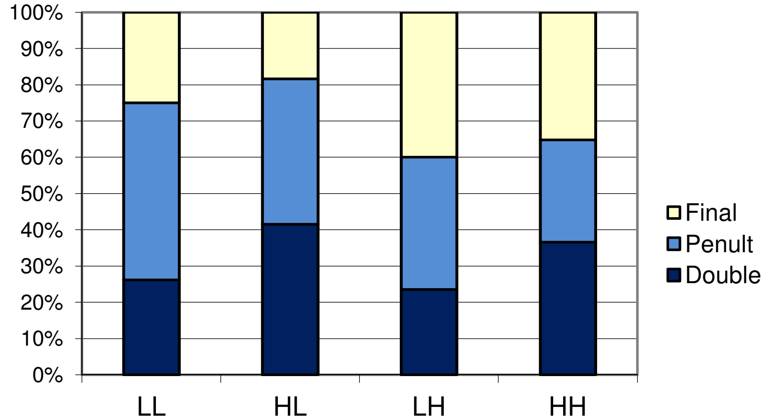
Table 9 Accent distribution of bisyllabic words according to syllable structure combinations

(The number given in parentheses is the percentage of each count)

Word Type	Accent Type			Total
	Double (%)	Penult (%)	Final (%)	
LL	65 (26)	**121 (49)**	62 (25)	248
HL	**86 (42)**	83 (40)	38 (18)	207
LH	63 (23)	98 (37)	**107 (40)**	268
HH	**52 (37)**	40 (28)	50 (35)	142
Total	266 (30)	**342 (40)**	257 (30)	865

The bolded numbers indicate the most frequent accent type for each combination. The table shows that the accent type with the highest frequency varies according to syllable structure: penultimate accent is most frequent (49%) in LL; double accent (42%) in HL; final accent (40%) in LH; and double accent in HH (37%). The proportional distribution of the different accent types is represented graphically in Figure 20:

Figure 20 Accent distribution in bisyllabic words according to syllable weight combinations



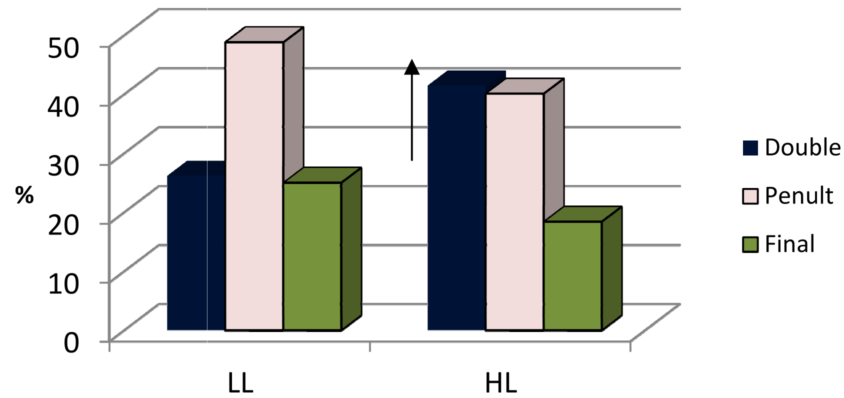
The data clearly demonstrate that the distribution patterns vary according to syllable structure combinations. In particular, we find the following generalizations:

(29) Generalizations about accent patterns in bisyllabic existing words

- penultimate accent is more frequent in LL
- double accent is more frequent and penultimate accent less frequent in HL, compared to LL and in LH
- final accent is more frequent and penultimate/double accent is less frequent in LH compared to LL and HL
- penultimate accent is less common in HH than in other combinations

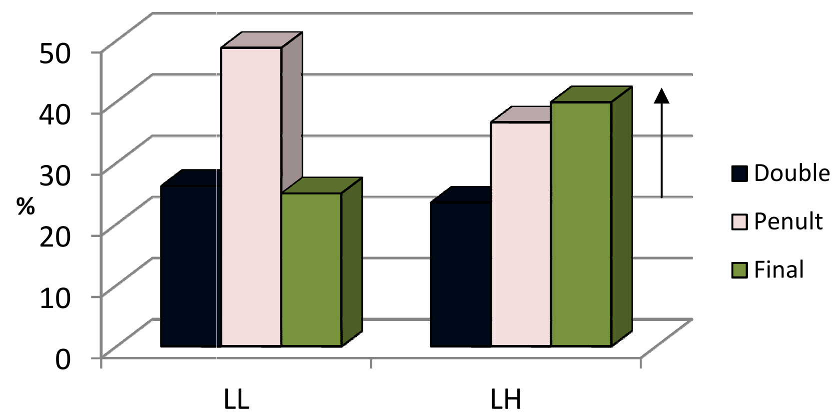
Statistical analyses were performed using Pearson's chi-square. The independent variables were syllable structure combination and accent position and the dependent variable was the number of tokens. In order to test for syllable structure effects on accent patterns, accentuation of other combinations (HL, LH, HH) was compared to accent in LL. In comparing accent distribution of HL with LL, double accent was found in 42% of HL forms but only in 26% of LL forms, which was significantly different ($\chi^2(2)$=12.16, p=.002).

Figure 21 Comparison of accent distribution in LL and in HL



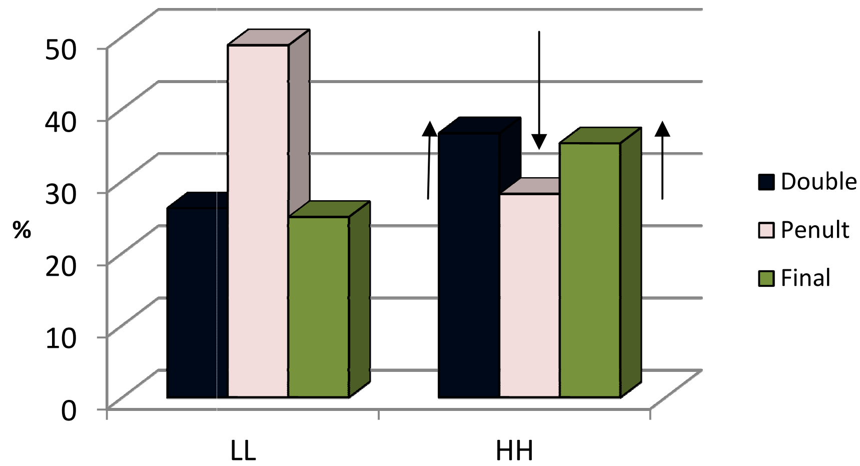
The accent distribution of LH was also significantly different from LL ($\chi^2(2)$=13.67, p=.001). Final accent appeared in 40% of LH forms, which was significantly more frequent than the 25% in LL. Penultimate accent was significantly less common (37%) in LH compared to 49% in LL.

Figure 22 Comparison of accent distribution in LL and in LH



HH combinations also showed a different accent distribution from LL forms ($\chi^2(2)$=15.84, p<.001). In accordance with the tendency for heavy syllables to attract accent, double accent was found in 37% of HH words vs. 26% of LL words. Furthermore, final accent was more frequent (35%) in HH than in LL (25%) due to a final heavy syllable. By contrast, penultimate accent was found in 28% of the forms, far less than the 49% in LL.

Figure 23 Comparison of accent distribution in LL and in HH

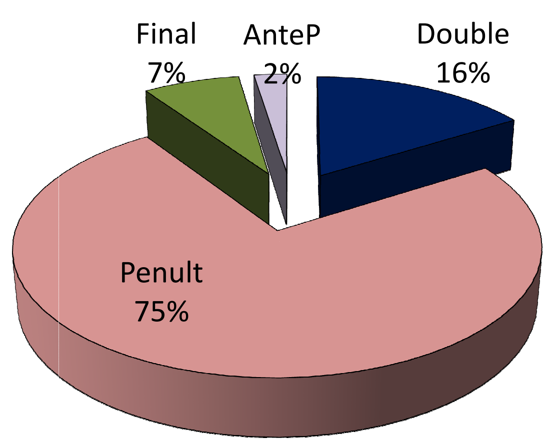


In addition, the distribution for HH forms is significantly different from HL ($\chi^2(2)=13.4$, $p=.001$) because of the lower frequency of penultimate accent (28% vs. 40%) and the higher frequency of final accent (35% vs. 18%). In sum, the accent distribution of bisyllabic words clearly reveals a preference for heavy syllables to be accented.

4.3.2 Trisyllabic words

Unlike the overall even distribution in bisyllabic words presented in Figure 19, trisyllabic words showed accent distribution skewed to penultimate accent: 75% of trisyllabic words (306 out of 411 words) had penultimate accent. Among the remaining 25%, 66 words (16%) had double accent, 29 words (7%) were given final accent, and only 10 words (2%) appeared with antepenultimate accent, as shown in Figure 24.

Figure 24 Accent distribution in trisyllabic words



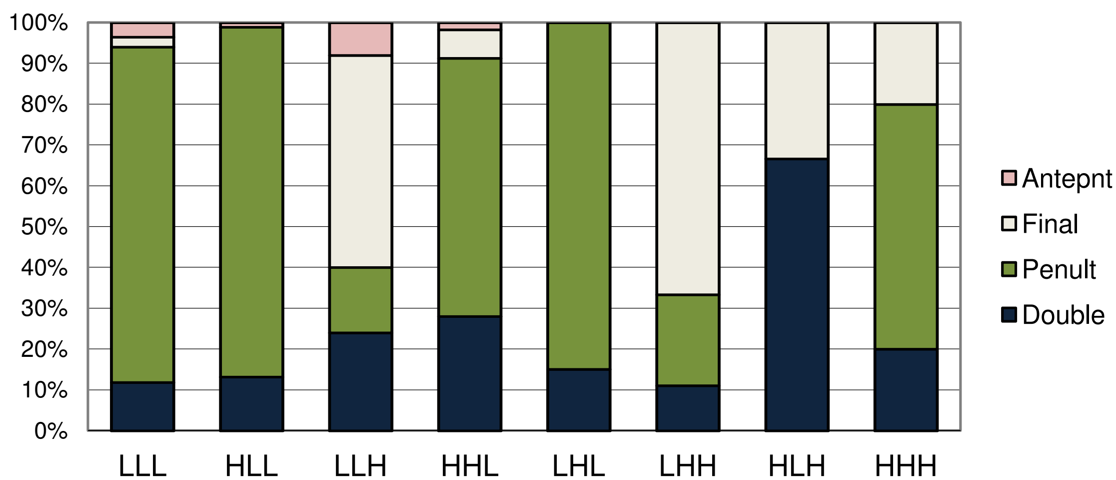
The distribution of accent in trisyllabic words according to syllable structure is presented in Table 10 and illustrated in Figure 25. Bolded numbers in the table indicate the most frequent accent for each combination.

Table 10 Accent distribution of trisyllabic words according to syllable structure

Word Type	Accent Type				Total
	Double (%)	Penult (%)	Final (%)	Antepenult[15] (%)	
LLL	20 (12)	**138 (82)**	4 (2)	6 (4)	168
HLL	12 (13)	**78 (86)**	0	1 (1)	91
LLH	6 (24)	4 (16)	**13 (52)**	2 (8)	25
HHL	16 (28)	**36 (63)**	4 (7)	1 (2)	57
LHL	8 (15)	**45 (85)**	0	0	53
LHH	1 (11)	2 (22)	**6 (67)**	0	9
HLH	**2 (67)**	0	1 (33)	0	3
HHH	1 (20)	**3 (60)**	1 (20)	0	5
Total	66 (16)	**306 (75)**	29 (7)	10 (2)	411

As shown in Table 10, penultimate accent is the most frequent in all combinations except LLH, LHH, and HLH: 82% in LLL; 86% in HLL; 63% in HHL; 85% in LHL; 60% in HHH. However, despite the overall predominance of penultimate accent, the preference for penultimate accent noticeably decreases in LLH, LHH, and HLH, where the final syllable is heavy, as illustrated in the bar graph in Figure 25.

Figure 25 Accent distribution of trisyllabic words according to syllable weight combinations



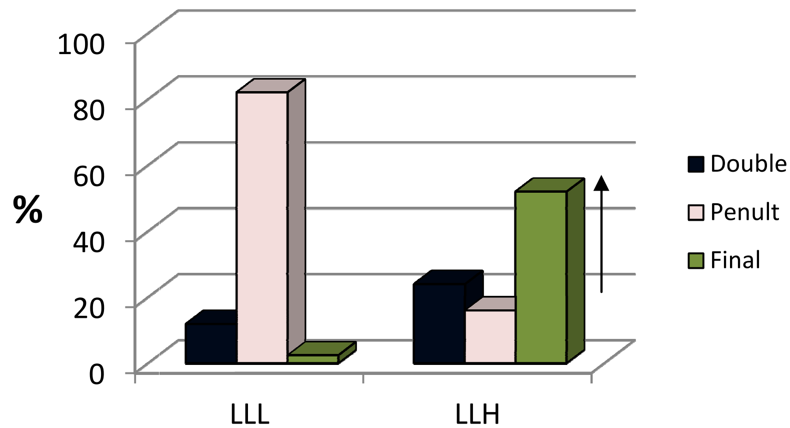
Penultimate accent is considerably less frequent in LLH, LHH and HLH compared to other combinations. In contrast, final accent is more frequent in the combinations which have a final heavy syllable. The association of HLH with more frequent double accent might not be

[15] Antepenult accent can be considered as initial accent since accent falls on the word-initial syllable.

representative of the lexical patterns because it was based on only 3 items, due to the rarity of the HLH combinations in the lexicon.

Pearson's chi-square tests were performed to see whether the difference in the lexical statistics was statistically significant. As shown in Figure 26, the test confirmed that the accent distribution in LLH vs. LLL were significantly different ($\chi^2(3)=77.16$, $p<.001$).

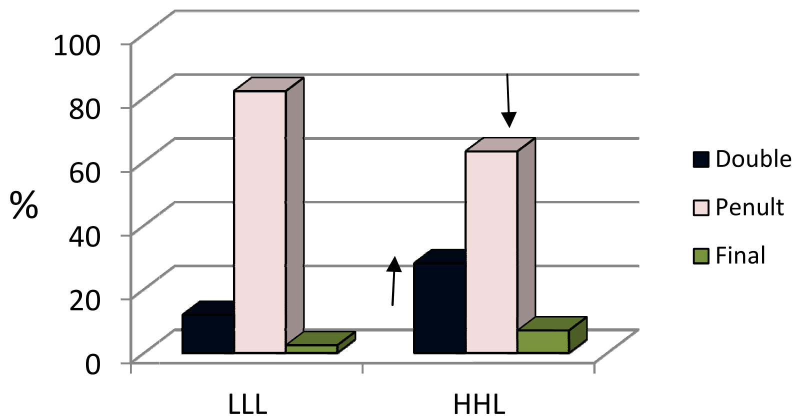
Figure 26 Comparison of accent distribution in LLL and in LLH



Penultimate accent is most frequent in LLL (82% vs. 16% in LLH) while final accent is most frequent in LLH (52% vs. 2% in LLL). This demonstrates that final accent assignment is strongly associated with a final heavy syllable and reveals a statistical tendency for heavy syllables to be accented.

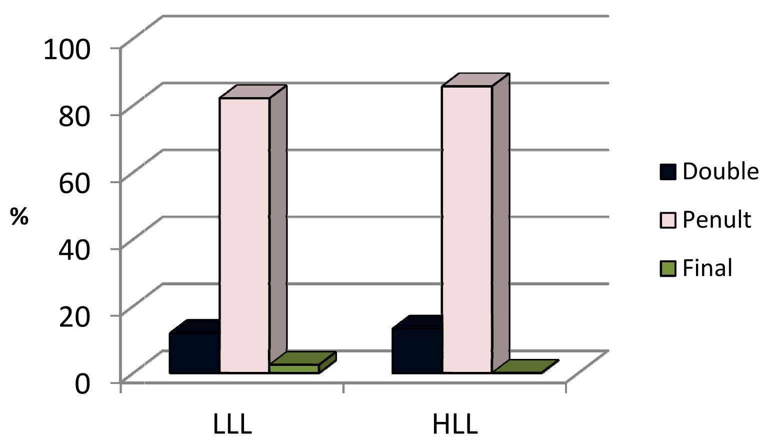
The syllable structure effects were also significant in the distribution of HHL vs. LLL ($\chi^2(3)=11.96$, $p=.008$), as shown in Figure 27. Even though penultimate accent is most common in HHL, as in LLL, double accent is more frequent in HHL than in LLL (28% vs. 12%) and penultimate accent is less frequent in HHL than in LLL (63% vs. 82%). Double accent in HHL is in accord with the tendency to accenting heavy syllables, since it results in no unaccented heavy syllables.

Figure 27 Comparison of accent distribution in LLL and in HHL



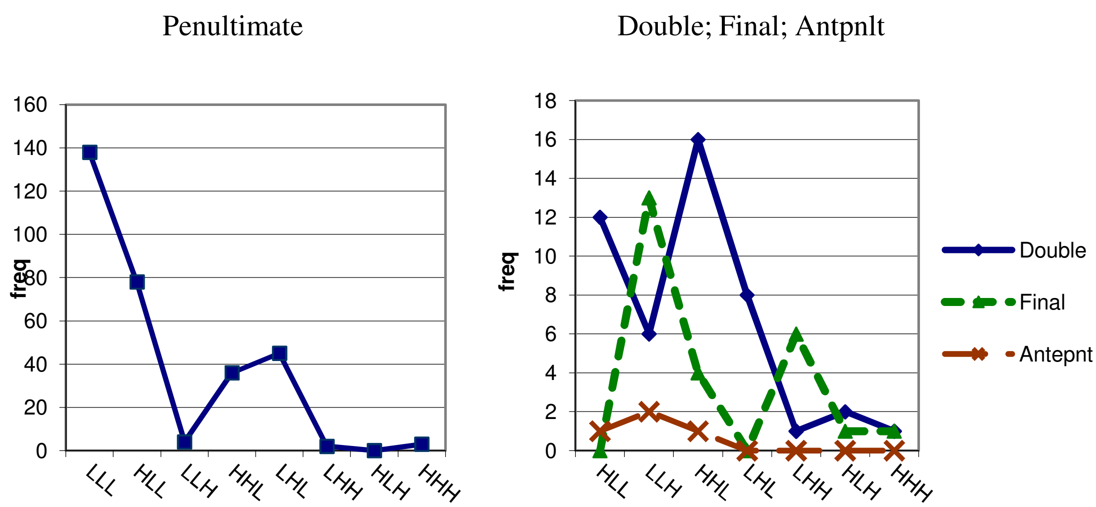
In comparing accent distribution of HLL with LLL, as shown in Figure 28, the accent pattern of HLL was not much different from LLL ($\chi^2(3)=3.67$, p=.3): in both combinations (HLL; LLL), penultimate accent is predominant (86%; 82%) and double accent is less common (13%; 12%). Double accent in HLL words was significantly fewer than in HHL words (13% vs. 28%, $\chi^2(3)=12.92$, p=.005). The statistical patterns of HLL words suggest that these words are more likely to exhibit the dominant penultimate accent rather than the preference for accent on a heavy syllable.

Figure 28 Comparison of accent distribution in LLL and in HLL



The overall prevalence of penultimate accent (75%) in trisyllabic words is not surprising, taking the skewed lexical distribution of the syllable structure combinations into consideration: about 70% of trisyllabic words were LLL, HHL, LHL combinations (40% were LLL combinations and 30% were HHL and LHL combinations). For those combinations, penultimate accent is not actually conflicting with the structure-sensitive tendencies. In addition, the following graphs in Figure 29 based on the raw frequency of the results help us to envision the input frequency of each accent that NKK learners are exposed to. The left-hand graph illustrates the occurrence of penultimate accent and the right-hand graph that of other accents.

Figure 29 Frequency of accent types in trisyllabic words

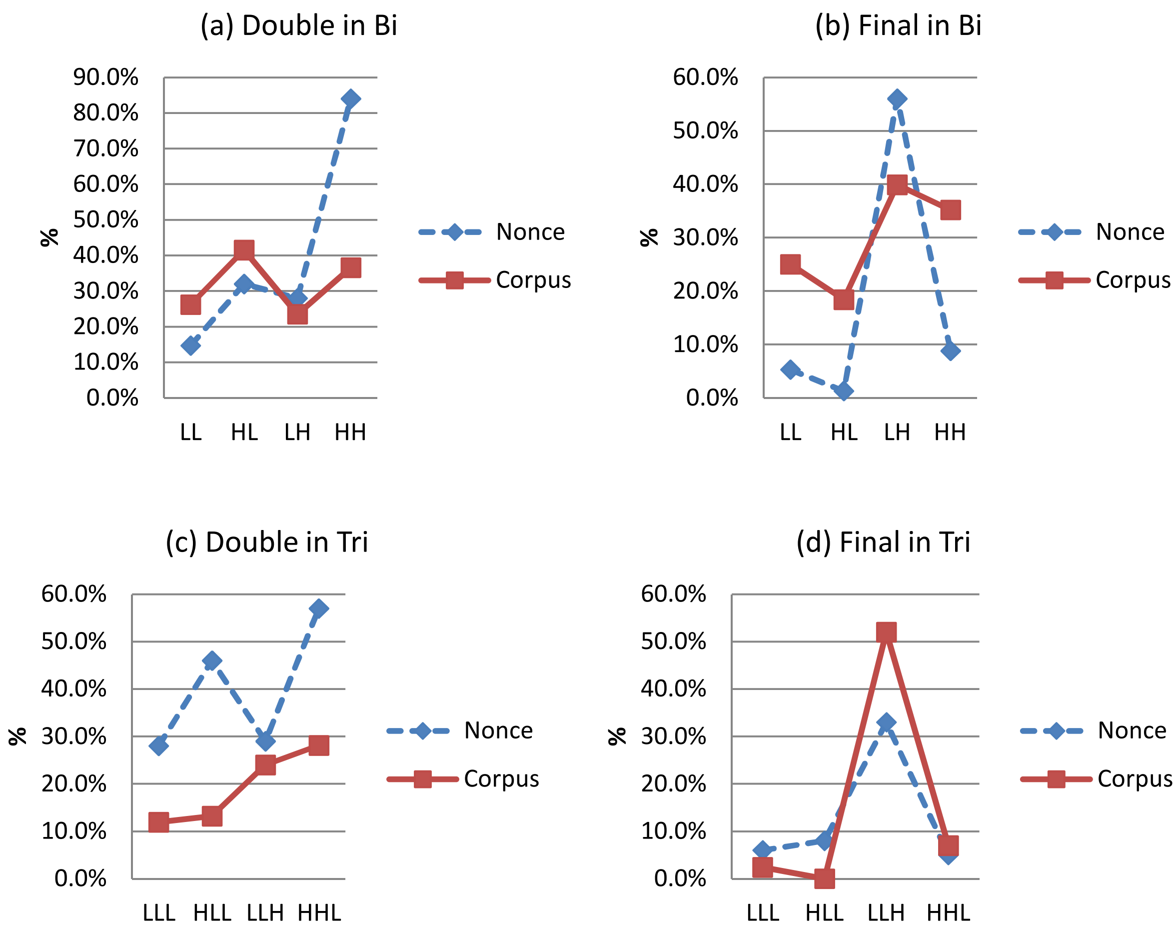


While penultimate accent was overwhelmingly frequent in trisyllabic words, this pattern is found with great frequency in LLL and HLL words, but far less frequently in other combinations. As illustrated in the right-hand graph in Figure 29, double accent was most frequent in HHL and final accent in LLH although double accent and final accent were not common in trisyllabic words. Therefore, the apparent patterns in the lexical statistics could be informative enough for NKK learners to sense subterranean structure-sensitive patterns in the lexicon and internalize them.

4.4 Lexical Frequency vs. Novel Words

The comparison of accent patterns in the corpus and in novel words revealed that the structure-sensitive accent tendencies emerging in novel words were indeed present in the statistical patterning in the lexicon, supporting the *Stochastic Accent Hypothesis*. The graphs in Figure 30 compare the accent patterns in the corpus and in novel words:

Figure 30 Comparison of accent patterns in the corpus and in novel words

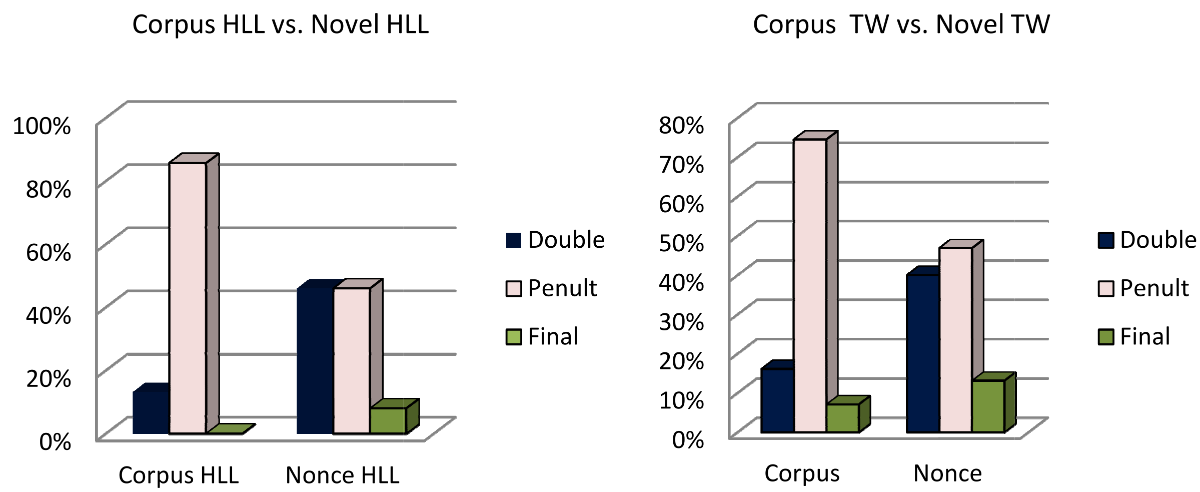


*Bi: bisyllabic words; Tri: trisyllabic words

As the graph (a) shows, double accent was most frequent in HL and HH bisyllabic words in the corpus, and novel words showed similar patterns for double accent. However, double accent in HH was far more frequent in novel words than in existing words. As shown in the graph (b), final accent was most frequent in LH both in the corpus and in novel words, but the tendency toward final accent was even stronger in novel words. Final accent in trisyllabic words, as shown in the graph (d), was rare in LLL, HLL, and HHL but common in LLH, both in novel words and in the corpus. The strong tendency toward accented word-final heavy syllables in novel words accords with the patterns revealed in the lexicon. The graph (c) illustrates that double accent was preferred in HHL both in novel words and in the corpus. The preference for double accent for HHL was even stronger in novel words. On the other hand, double accent in HLL was not as common as in HHL in the corpus, whereas double accent was stronger in HLL novel words.

The statistical tendencies in the corpus were consistent with the syllable structure-sensitive patterning in novel words in general. However, for double accent assignment, the quantitative difference in novel words vs. in the corpus was noticeable, as shown in the graph (c) in Figure 30. For example, in HLL combinations, the difference in double accent assignment between novel words and the corpus is considerable (46% in novel vs. 13% in corpus), resulting in distinct accent distribution for the two groups: the predominance of penultimate accent in the corpus HLL words contrasts with the weaker tendency toward penultimate accent in the novel HLL forms, as illustrated in the left-hand graph in Figure 31. The strong preference for double accent in the novel HLL forms, despite the low frequency of double accent in the corresponding syllable structure combinations in the lexicon, can be accounted for if we assume that structure-sensitivity outweighs lexical frequency, with NKK speakers preferring the less frequent structure-sensitive patterning to patterns which contradict the structure-sensitive patterning, even when the structure insensitive patterns are more frequent.

Figure 31 Accent distribution in trisyllabic words: Corpus vs. Novel words



In consequence, the overall accent distribution in the corpus differed from that in novel words, as shown in the right-hand graph in Figure 31: the overall prevalence of penultimate accent in trisyllabic words in the corpus (75%) was not reflected in the novel words, where double accent was assigned more frequently.

To summarize, the structure-sensitive accent tendencies emerging in novel words were generally consistent with the statistical tendencies in the corpus. This suggests that NKK speakers indeed have an implicit knowledge of the structure-sensitive patterning reflected in the lexicon and extend it to novel words. However, NKK speakers do not simply replicate the lexical frequency of accent patterns since the structure-sensitivity in accentuation was even stronger in novel words than in the corpus.

4.5 Conclusion

Previous studies of NKK accentuation have assumed that NKK lexical accent patterns should not be associated with syllable structure because syllable structure does not predict lexical accent patterns categorically (G-R. Kim 1988; Y-H. Chung 1991; N-J. Kim 1997; S-H. Kim 1999). However, the corpus study revealed that syllable structure indeed matters in the probabilistic accent tendencies in the lexicon: there is a strong tendency toward final accent on LH and LLH; a tendency toward double accent on HH and HHL; and a tendency toward penultimate accent elsewhere. The statistical accent patterns in the lexicon were generally consistent with the structure-sensitive accent tendencies emerging in novel words, which strongly suggest that NKK speakers indeed internalize the statistical patterning and apply it to novel words. Furthermore, structure-sensitivity was even stronger in novel words than in the corpus, which indicates that NKK speakers do not simply generalize the lexical statistical tendencies but reproduce phonologically constrained patterning in novel words. Thus, the results support *the Stochastic Accent Hypothesis* and a stochastic grammar based on lexical patterns.

This conclusion is consistent with patterns in existing loanwords from English: penultimate accent is assigned in words without heavy syllables (e.g. *ʃik^{h}áko* ‘Chicago’), final accent in words ending in a heavy syllable (e.g. *sit^{h}ak^{h}íŋ* 'stocking'), and double accent in words beginning with heavy syllables (e.g. *símp^{h}óni* ‘symphony’) (Kenstowicz and Sohn 2001; Y-H Chung 2002, 2006). The structure-sensitive patterns of loanwords cannot be attributed to the influence of the source language because accent placement in loanwords is not necessarily consistent with stress placement in the source language (e.g. *sit^{h}ak^{h}íŋ* vs. stócking). Rather, the accent patterns in loanwords are in accordance with the native language association between syllable weight and accent. This implies that the syllable structure effects emerging in on-line accent assignment in newly adopted words (novel words) persist in off-line accent assignment in long-lasting loanwords.

Appendix II: Accent Patterns in the Corpus

NO	Item	Transcription*	No.of Syllables	Syll.Str. Combination	Accent Type**		
				L:CV, H:CVC	NKK1	NKK2	NKK3
1	얘기	jeki	2	LL	D	D	D
2	아빠	appa	2	LL	D	D	D
3	나이	nai	2	LL	D	D	D
4	아기	aki	2	LL	D	D	D
5	오빠	oppa	2	LL	D	D	D
6	가게	kake	2	LL	D	D	D
7	거리	kəri	2	LL	D	D	D
8	아무	amu	2	LL	D	D	D
9	어미	əmi	2	LL	D	D	D
10	바보	papo	2	LL	D	D	D
11	추위	chuwi	2	LL	D	D	D
12	소매	some	2	LL	D	D	D
13	배추	pechu	2	LL	D	D	D
14	토끼	thokki	2	LL	D	D	D
15	조카	cokha	2	LL	D	D	D
16	대꾸	tekku	2	LL	D	D	D
17	파리	phari	2	LL	D	D	D

18	겨레	kjəre	2	LL	D	D	D
19	겨자	kjəca	2	LL	D	D	P
20	까치	kkachi	2	LL	D	D	D
21	수고	suko	2	LL	D	D	D
22	모기	moki	2	LL	D	D	D
23	대추	techu	2	LL	D	D	D
24	오리	ori	2	LL	D	D	D
25	나귀	nakwi	2	LL	D	D	D
26	데뷔	tepwi	2	LL	D	D	D
27	수다	suta	2	LL	D	D	D
28	주제	cuce	2	LL	D	D	D
29	베개	peke	2	LL	D	D	D
30	채비	chepi	2	LL	D	D	F
31	표고	phjoko	2	LL	D	D	D
32	도끼	tokki	2	LL	D	D	D
33	모레	more	2	LL	D	D	D
34	시내	sine	2	LL	D	D	D
35	나리	nari	2	LL	D	D	P
36	해태	hethe	2	LL	D	D	D
37	부추	puchu	2	LL	D	D	D
38	가재	kace	2	LL	D	D	D
39	가지 1(‘branch’)	kaci	2	LL	D	D	F
40	대패	tephe	2	LL	D	D	D
41	이리	iri	2	LL	D	D	D
42	가래	kare	2	LL	D	D	D
43	부아	pua	2	LL	D	D	D
44	뽀뽀	ppoppo	2	LL	D	D	D
45	아가	aka	2	LL	D	D	D
46	무쇠	muswe	2	LL	D	D	D
47	네모	nemo	2	LL	D	D	D
48	보배	pope	2	LL	D	D	D
49	차려	charjə	2	LL	D	D	D
50	투구	thuku	2	LL	D	D	D
51	수구	suku	2	LL	D	D	D
52	시위	siwi	2	LL	D	D	D
53	어치	əchi	2	LL	D	D	P
54	허파	həpha	2	LL	D	D	D
55	보도	poto	2	LL	D	D	D
56	세모	semo	2	LL	D	D	D
57	제기	ceki	2	LL	D	D	D
58	고지	koci	2	LL	D	D	P

59	부도	puto	2	LL	D	D	D
60	여치	jechi	2	LL	D	D	D
61	오뉘	onwi	2	LL	D	D	F
62	자궤	cakwe	2	LL	D	D	D
63	주체	cuche	2	LL	D	D	D
64	쥐치	cwichi	2	LL	D	D	D
65	하마	hama	2	LL	D	D	D
66	하나	hana	2	LL	F	F	F
67	하루	haru	2	LL	F	F	F
68	나무	namu	2	LL	F	F	F
69	다리	tari	2	LL	F	F	F
70	사내	sane	2	LL	F	F	F
71	재미	cemi	2	LL	F	F	F
72	서로	səro	2	LL	F	F	F
73	가지 3(‘eggplant’)	kaci	2	LL	F	F	F
74	고추	kochu	2	LL	F	F	F
75	후추	huchu	2	LL	F	F	D
76	마루	maru	2	LL	F	F	F
77	치마	chima	2	LL	F	F	F
78	가루	karu	2	LL	F	F	F
79	마디	mati	2	LL	F	F	F
80	구두	kutu	2	LL	F	F	F
81	꼬마	kkoma	2	LL	F	F	F
82	보리	pori	2	LL	F	F	F
83	수수	susu	2	LL	F	F	F
84	여우	jəu	2	LL	F	F	F
85	조개	coke	2	LL	F	F	F
86	처마	chəma	2	LL	F	F	F
87	고무	komu	2	LL	F	F	F
88	모시	mosi	2	LL	F	F	F
89	오이	oi	2	LL	F	F	F
90	아우	au	2	LL	F	F	F
91	호두	hotu	2	LL	F	F	F
92	가로	karo	2	LL	F	F	F
93	지네	cine	2	LL	F	F	F
94	고래	kore	2	LL	F	F	F
95	기와	kiwa	2	LL	F	F	F
96	나루	naru	2	LL	F	F	F
97	세로	sero	2	LL	F	F	F
98	시루	siru	2	LL	F	F	F
99	노루	noru	2	LL	F	F	F

100	자루	caru	2	LL	F	F	F
101	보라	pora	2	LL	F	F	F
102	두메	tume	2	LL	F	F	F
103	추녀	chunjə	2	LL	F	F	D
104	고요	kojo	2	LL	F	F	F
105	어이	əi	2	LL	F	F	F
106	그네	kɨne	2	LL	F	F	F
107	모루	moru	2	LL	F	F	F
108	소라	sora	2	LL	F	F	F
109	배래	pere	2	LL	F	F	?
110	지레	cire	2	LL	F	F	F
111	두부	tupu	2	LL	F	F	F
112	그루	kɨru	2	LL	F	F	P
113	다래	tare	2	LL	F	F	F
114	부처	puchə	2	LL	F	F	F
115	비녀	pinjə	2	LL	F	F	F
116	자라	cara	2	LL	F	F	F
117	두레	ture	2	LL	F	F	F
118	우레	ure	2	LL	F	F	F
119	자두	catu	2	LL	F	F	F
120	푸대	phute	2	LL	F	F	F
121	뒤주	twicu	2	LL	F	F	F
122	머루	məru	2	LL	F	F	F
123	모이	moi	2	LL	F	F	P
124	사래	sare	2	LL	F	F	F
125	소태	sothe	2	LL	F	F	F
126	아구	aku	2	LL	F	F	F
127	자개	cake	2	LL	F	F	F
128	소리	sori	2	LL	P	P	P
129	아이	ai	2	LL	P	P	P
130	사이	sai	2	LL	P	P	P
131	나라	nara	2	LL	P	P	P
132	자리	cari	2	LL	P	P	P
133	머리	məri	2	LL	P	P	P
134	고개	koke	2	LL	P	P	P
135	아내	ane	2	LL	P	P	P
136	아래	are	2	LL	P	P	P
137	노래	nore	2	LL	P	P	P
138	바다	pata	2	LL	P	P	P
139	모두	motu	2	LL	P	P	P
140	어깨	əkke	2	LL	P	P	P

141	고기	koki	2	LL	P	P	P
142	허리	həri	2	LL	P	P	P
143	누나	nuna	2	LL	P	P	P
144	뿌리	ppuri	2	LL	P	P	P
145	새끼	sekki	2	LL	P	P	P
146	꼬리	kkori	2	LL	P	P	P
147	너머	nəmə	2	LL	P	P	P
148	어제	əce	2	LL	P	P	P
149	바위	pawi	2	LL	P	P	P
150	무게	muke	2	LL	P	P	P
151	이마	ima	2	LL	P	P	P
152	바지	paci	2	LL	P	P	P
153	모래	more	2	LL	P	P	P
154	돼지	tweci	2	LL	P	P	P
155	더위	təwi	2	LL	P	P	P
156	무늬	muni	2	LL	P	P	P
157	무리	muri	2	LL	P	P	P
158	개미	kemi	2	LL	P	P	P
159	제비	cepi	2	LL	P	P	P
160	아씨	assi	2	LL	P	P	P
161	또래	ttore	2	LL	P	P	P
162	사위	sawi	2	LL	P	P	P
163	두께	tukke	2	LL	P	P	P
164	거지	kəci	2	LL	P	P	P
165	새우	seu	2	LL	P	P	P
166	재주	cecu	2	LL	P	P	P
167	바퀴	pakhwi	2	LL	P	P	P
168	고리	kori	2	LL	P	P	P
169	아비	api	2	LL	P	P	P
170	나비	napi	2	LL	P	P	P
171	가위	kawi	2	LL	P	P	P
172	더미	təmi	2	LL	P	P	P
173	누이	nui	2	LL	P	P	P
174	비누	pinu	2	LL	P	P	P
175	찌개	ccike	2	LL	P	P	P
176	개비	kepi	2	LL	P	P	P
177	자취	cachwi	2	LL	P	P	D
178	거미	kəmi	2	LL	P	P	P
179	부리	puri	2	LL	P	P	P
180	부채	puche	2	LL	P	P	P
181	도마	toma	2	LL	P	P	P

182	수레	sure	2	LL	P	P	F
183	어귀	əkwi	2	LL	P	P	P
184	미끼	mikki	2	LL	P	P	P
185	수저	sucə	2	LL	P	P	P
186	지게	cike	2	LL	P	P	P
187	고삐	koppi	2	LL	P	P	P
188	이끼	ikki	2	LL	P	P	P
189	포기(‘head’)	phoki	2	LL	P	P	P
190	대야	teja	2	LL	P	P	P
191	부피	puphi	2	LL	P	P	P
192	가시	kasi	2	LL	P	P	P
193	가마 (‘kiln’)	kama	2	LL	P	P	P
194	보기	poki	2	LL	P	P	P
195	서리	səri	2	LL	P	P	P
196	뙈기	ttweki	2	LL	P	P	P
197	메주	mecu	2	LL	P	P	P
198	꼬치	kkochi	2	LL	P	P	P
199	너비	nəpi	2	LL	P	P	P
200	구리	kuri	2	LL	P	P	P
201	매미	memi	2	LL	P	P	P
202	가지 2 (‘kind’)	kaci	2	LL	P	P	P
203	끼니	kkini	2	LL	P	P	P
204	쐐기	ssweki	2	LL	P	P	P
205	누에	nue	2	LL	P	P	P
206	맏이	maci	2	LL	P	P	P
207	이모	imo	2	LL	P	P	P
208	피리	phiri	2	LL	P	P	P
209	벼루	pjəru	2	LL	P	P	P
210	따귀	ttakwi	2	LL	P	P	D
211	마개	make	2	LL	P	P	P
212	도미	tomi	2	LL	P	P	P
213	메기	meki	2	LL	P	P	D
214	호미	homi	2	LL	P	P	P
215	기미	kimi	2	LL	P	P	P
216	도리	tori	2	LL	P	P	D
217	비계	pikje	2	LL	P	P	P
218	새치	sechi	2	LL	P	P	D
219	싸리	ssari	2	LL	P	P	?
220	아귀	akwi	2	LL	P	P	P
221	아퀴	akhwi	2	LL	P	P	P
222	조리	cori	2	LL	P	P	P

223	그제	kɨce	2	LL	P	P	F
224	나래	nare	2	LL	P	P	P
225	누비	nupi	2	LL	P	P	P
226	따비	ttapi	2	LL	P	P	?
227	세무	semu	2	LL	P	P	D
228	애꾸	ekku	2	LL	P	P	P
229	거위	kəwi	2	LL	P	P	P
230	고비	kopi	2	LL	P	P	P
231	꽈리	kkwari	2	LL	P	P	P
232	누리	nuri	2	LL	P	P	P
233	똬리	ttwari	2	LL	P	P	P
234	머위	məwi	2	LL	P	P	?
235	배미	pemi	2	LL	P	P	P
236	버찌	pəcci	2	LL	P	P	P
237	버캐	pəkhe	2	LL	P	P	?
238	비루	piru	2	LL	P	P	P
239	비지	pici	2	LL	P	P	P
240	소개	soke	2	LL	P	P	P
241	아재	ace	2	LL	P	P	P
242	에미	emi	2	LL	P	P	P
243	우세	use	2	LL	P	P	P
244	자구	caku	2	LL	P	P	?
245	자귀	cakwi	2	LL	P	P	P
246	자위 (‘iris’)	cawi	2	LL	P	P	P/D
247	태주	thecu	2	LL	P	P	?
248	엄마	əmma	2	HL	D	D	D
249	얼마	əlma	2	HL	D	D	D
250	담배	tampe	2	HL	D	D	D
251	냄새	nemse	2	HL	D	D	D
252	날씨	nalssi	2	HL	D	D	D
253	날개	nalke	2	HL	D	D	D
254	장사	caŋsa	2	HL	D	D	D
255	낚시	naksi	2	HL	D	D	D
256	솜씨	somssi	2	HL	D	D	D
257	날짜	nalcca	2	HL	D	D	D
258	일쑤	ilssu	2	HL	D	D	D
259	잔치	canchi	2	HL	D	D	D
260	안개	aŋke	2	HL	D	D	D
261	글씨	kɨlssi	2	HL	D	D	D
262	열쇠	jəlswe	2	HL	D	D	D
263	늑대	nɨkte	2	HL	D	D	D

264	참외	chamwe	2	HL	D	D	D
265	번개	pəŋke	2	HL	D	D	D
266	할매	halme	2	HL	D	D	D
267	골치	kolchi	2	HL	D	D	D
268	방귀	paŋkwi	2	HL	D	D	D
269	장가	caŋka	2	HL	D	D	D
270	엄두	əmtu	2	HL	D	D	D
271	억지	əkci	2	HL	D	D	D
272	전기	cəŋki	2	HL	D	D	D
273	상투	saŋthu	2	HL	D	D	D
274	딸기	ttalki	2	HL	D	D	D
275	임자	imca	2	HL	D	D	D
276	담요	tamjo	2	HL	D	D	D
277	돌기	tolki	2	HL	D	D	D
278	붕어	puŋə	2	HL	D	D	D
279	염소	jəmso	2	HL	D	D	D
280	색시	seksi	2	HL	D	D	D
281	인도	into	2	HL	D	D	D
282	잉어	iŋə	2	HL	D	D	D
283	톱니	thomni	2	HL	D	D	D
284	광대	kwaŋte	2	HL	D	D	D
285	쓸모	ssɨlmo	2	HL	D	D	D
286	갈대	kalte	2	HL	D	D	D
287	장끼	caŋkki	2	HL	D	D	D
288	난로	nallo	2	HL	D	D	D
289	억새	əkse	2	HL	D	D	D
290	박새	pakse	2	HL	D	D	D
291	올케	olkhe	2	HL	D	D	D
292	물꼬	mulkko	2	HL	D	D	D
293	진지	cinci	2	HL	D	D	D
294	박쥐	pakcwi	2	HL	D	D	D
295	갈치	kalchi	2	HL	D	D	D
296	꼴찌	kkocci	2	HL	D	D	D
297	낙지	nakci	2	HL	D	D	D
298	쓸개	ssɨlke	2	HL	D	D	D
299	솔기	solki	2	HL	D	D	D
300	창자	chaŋca	2	HL	D	D	D
301	냉이	neŋi	2	HL	D	D	D
302	단지	tanci	2	HL	D	D	D
303	삼베	sampe	2	HL	D	D	D

304	숙주	sukcu	2	HL	D	D	D
305	숭어	suŋə	2	HL	D	D	D
306	억수	əksu	2	HL	D	D	D
307	곰보	kompo	2	HL	D	D	D
308	꽁치	kkoŋchi	2	HL	D	D	D
309	녹초	nokcho	2	HL	D	D	D
310	망새	maŋse	2	HL	D	D	D
311	맹추	meŋchu	2	HL	D	D	D
312	볼모	polmo	2	HL	D	D	D
313	푼수	phunsu	2	HL	D	D	D
314	곤로	kollo	2	HL	D	D	?
315	높새	nopse	2	HL	D	D	D
316	덧니	tənni	2	HL	D	D	D
317	망개	maŋke	2	HL	D	D	D
318	멍게	məŋke	2	HL	D	D	D
319	몰매	molme	2	HL	D	D	D
320	밀대	milte	2	HL	D	D	D
321	벅수	pəksu	2	HL	D	D	D
322	완자	wanca	2	HL	D	D	D
323	퉁소	thuŋso	2	HL	D	D	P
324	동티	toŋthi	2	HL	D	D	?
325	들보	tɨlpo	2	HL	D	D	D
326	목로	moŋno	2	HL	D	D	D
327	방게	paŋke	2	HL	D	D	D
328	성게	səŋke	2	HL	D	D	D
329	숙수	suksu	2	HL	D	D	D
330	알짜	alcca	2	HL	D	D	D
331	양치	jaŋchi	2	HL	D	D	D
332	율무	julmu	2	HL	D	D	D
333	빨래	ppalle	2	HL	F	F	F
334	먼지	mənci	2	HL	F	F	F
335	감자	kamca	2	HL	F	F	F
336	열매	jəlme	2	HL	F	F	F
337	막내	maŋne	2	HL	F	F	F
338	냄비	nempi	2	HL	F	F	F
339	잔디	canti	2	HL	F	F	F
340	장마	caŋma	2	HL	F	F	F
341	엄지	əmci	2	HL	F	F	F
342	인제	ince	2	HL	F	F	F
343	엄포	əmpho	2	HL	F	F	F

344	단추	tanchu	2	HL	F	F	F
345	감투	kamthu	2	HL	F	F	F
346	빈대	pinte	2	HL	F	F	F
347	썰매	ssəlme	2	HL	F	F	F
348	앵두	eŋtu	2	HL	F	F	F
349	연어	jənə	2	LL	F	F	F
350	찔레	ccille	2	HL	F	F	F
351	상어	saŋə	2	HL	F	F	F
352	연지	jənci	2	HL	F	F	F
353	열무	jəlmu	2	HL	F	F	F
354	꼽추	kkopchu	2	HL	F	F	F
355	달래	talle	2	HL	F	F	F
356	순대	sunte	2	HL	F	F	F
357	곤지	konci	2	HL	F	F	F
358	넝마	nəŋma	2	HL	F	P	F
359	명주	mjəŋcu	2	HL	F	F	F
360	방아	paŋa	2	HL	F	F	F
361	살구	salku	2	HL	F	F	F
362	상고	saŋko	2	HL	F	F	F
363	선지	sənci	2	HL	F	F	F
364	술래	sulle	2	HL	F	F	F
365	얼레	əlle	2	HL	F	F	F
366	웬수	wensu	2	HL	F	F	F
367	작두	caktu	2	HL	F	F	D
368	절구	cəlku	2	HL	F	F	F
369	정지	cəŋci	2	HL	F	F	F
370	혼자	honca	2	HL	P	P	P
371	언니	ənni	2	HL	P	P	P
372	종이	coŋi	2	HL	P	P	P
373	김치	kimchi	2	HL	P	P	P
374	줄기	culki	2	HL	P	P	P
375	선비	səmpi	2	HL	P	P	P
376	둥지	tuŋci	2	HL	P	P	P
377	흉내	hjuŋne	2	HL	P	P	P
378	벌레	pəlle	2	HL	P	P	P
379	접시	cəpsi	2	HL	P	P	P
380	신세	sinse	2	HL	P	P	P
381	핑계	phiŋke	2	HL	P	P	P
382	동무	toŋmu	2	HL	P	P	P
383	멸치	mjəchi	2	HL	P	P	P

384	덩치	təŋchi	2	HL	P	P	P
385	국수	kuksu	2	HL	P	P	P
386	송이	soŋi	2	HL	P	P	P
387	장수	caŋsu	2	HL	P	P	P
388	둘레	tulle	2	HL	P	P	P
389	갈래	kalle	2	HL	P	P	P
390	덩이	təŋi	2	HL	P	P	P
391	막대	makte	2	HL	P	P	P
392	슬기	sɨlki	2	HL	P	P	D
393	꼭지	kkokci	2	HL	P	P	P
394	상추	saŋchu	2	HL	P	P	P
395	참치	chamchi	2	HL	P	P	P
396	갈비	kalpi	2	HL	P	P	P
397	걸레	kəlle	2	HL	P	P	P
398	굴레	kulle	2	HL	P	P	P
399	꽁초	kkoŋcho	2	HL	P	P	P
400	뭉치	muŋchi	2	HL	P	P	P
401	팽이	pheŋi	2	HL	P	P	P
402	빌미	pilmi	2	HL	P	P	D
403	행주	heŋcu	2	HL	P	P	P
404	댕기	teŋki	2	HL	P	P	P
405	망치	maŋchi	2	HL	P	P	P
406	갈피	kalphi	2	HL	P	P	P
407	쟁기	ceŋki	2	HL	P	P	P
408	잎새	ipse	2	HL	P	P	D
409	각시	kaksi	2	HL	P	P	P
410	낌새	kkimse	2	HL	P	P	P
411	둥치	tuŋchi	2	HL	P	P	P
412	삽바	sapppa	2	HL	P	P	P
413	집게	cipke	2	HL	P	P	P
414	품바	phumpa	2	HL	P	P	P
415	괭이	kweŋi	2	HL	P	P	D
416	국자	kukca	2	HL	P	P	P
417	딱지	ttakci	2	HL	P	P	P
418	맵시	mepsi	2	HL	P	P	P
419	명치	mjəŋchi	2	HL	P	P	F
420	덜미	təlmi	2	HL	P	P	P
421	짱구	ccaŋku	2	HL	P	P	P
422	쪽지	ccokci	2	HL	P	P	P

423	흉터	hjuŋthə	2	HL	P	P	P
424	날치	nalchi	2	HL	P	P	P
425	편수	pjənsu	2	HL	P	P	P
426	공기	koŋki	2	HL	P	P	P
427	꺾쇠	kkəkswe	2	HL	P	P	P
428	말미	malmi	2	HL	P	P	D
429	망초	maŋcho	2	HL	P	P	P
430	멀미	məlmi	2	HL	P	P	P
431	멍에	məŋe	2	HL	P	P	P
432	싹수	ssaksu	2	HL	P	P	P
433	굴비	kulpi	2	HL	P	P	P
434	깍지	kkakci	2	HL	P	P	P
435	얌체	jamche	2	HL	P	P	P
436	장구	caŋku	2	HL	P	P	P
437	함지	hamci	2	HL	P	P	P
438	꽁지	kkoŋci	2	HL	P	P	P
439	뚱보	ttuŋpo	2	HL	P	P	P
440	볼기	polki	2	HL	P	P	D
441	핑계	phiŋkje	2	HL	P	P	P
442	갈퀴	kalkhwi	2	HL	P	P	P
443	강짜	kaŋcca	2	HL	P	P	P
444	글피	kɨlphi	2	HL	P	P	P
445	냄시	nemsi	2	HL	P	P	?
446	삿대	satte	2	HL	P	P	D
447	손주	soncu	2	HL	P	P	P
448	얼개	əlke	2	HL	P	P	P
449	종지	coŋci	2	HL	P	P	P
450	죽지	cukci	2	HL	P	P	P
451	탱자	theŋca	2	HL	P	P	P
452	흘레	hɨlle	2	HL	P	P	P
453	사람	saram	2	LH	D	D	D
454	처음	chəɨm	2	LH	D	D	D
455	그림	kɨrim	2	LH	D	D	D
456	어른	ərɨn	2	LH	D	D	D
457	나중	nacuŋ	2	LH	D	D	D
458	과일	kwail	2	LH	D	D	D
459	시골	sikol	2	LH	D	D	D
460	거짓	kəcit	2	LH	D	D	D
461	구름	kurɨm	2	LH	D	D	D

462	구경	kukjəŋ	2	LH	D	D	D
463	기운	kiun	2	LH	D	D	D
464	호박	hopak	2	LH	D	D	D
465	수박	supak	2	LH	D	D	D
466	벼락	pjərak	2	LH	D	D	D
467	무좀	mucom	2	LH	D	D	D
468	이빨	ippal	2	LH	D	D	D
469	서랍	sərap	2	LH	D	D	F
470	타령	tharjəŋ	2	LH	D	D	D
471	그믐	kɨmɨm	2	LH	D	D	D
472	저승	cəsɨŋ	2	LH	D	D	D
473	터전	thəcən	2	LH	D	D	D
474	그물	kɨmul	2	LH	D	D	D
475	시늉	sinjuŋ	2	LH	D	D	D
476	소름	sorɨm	2	LH	D	D	D
477	무당	mutaŋ	2	LH	D	D	D
478	푸념	phunjəm	2	LH	D	D	D
479	개펄	kephəl	2	LH	D	D	D
480	내숭	nesuŋ	2	LH	D	D	D
481	마님	manim	2	LH	D	D	D
482	오금	okɨm	2	LH	D	D	D
483	외동	wetoŋ	2	LH	D	D	D
484	주눅	cunuk	2	LH	D	D	D
485	세간	sekan	2	LH	D	D	D
486	트림	thɨrim	2	LH	D	D	D
487	과녁	kwanjək	2	LH	D	D	D
488	서울	səul	2	LH	D	D	D
489	트집	thɨcip	2	LH	D	D	D
490	귀양	kwijaŋ	2	LH	D	D	D
491	시중	sicuŋ	2	LH	D	D	D
492	새알	seal	2	LH	D	D	D
493	도령	torjəŋ	2	LH	D	D	D
494	때깔	ttekkal	2	LH	D	D	D
495	부축	puchuk	2	LH	D	D	D
496	이골	ikol	2	LH	D	D	D
497	마전	macən	2	LH	D	D	D
498	메밥	mepap	2	LH	D	D	D
499	소경	sokjəŋ	2	LH	D	D	D
500	시상	sisaŋ	2	LH	D	D	D

501	주접	cucəp	2	LH	D	D	D
502	타박	thapak	2	LH	D	D	D
503	가물	kamul	2	LH	D	D	D
504	무등	mutɨŋ	2	LH	D	D	D
505	미립	mirip	2	LH	D	D	D
506	배알	peal	2	LH	D	D	D
507	배웅	peuŋ	2	LH	D	D	D
508	부삽	pusap	2	LH	D	D	D
509	아둔	atun	2	LH	D	D	D
510	애살	esal	2	LH	D	D	D
511	자석	casək	2	LH	D	D	D
512	주낙	cunak	2	LH	D	D	D
513	주변	supjən	2	LH	D	D	D
514	마음	maɨm	2	LH	F	F	F
515	아침	achim	2	LH	F	F	F
516	사랑	saraŋ	2	LH	F	F	F
517	요즘	jocɨm	2	LH	F	F	F
518	바람	param	2	LH	F	F	F
519	마을	maɨl	2	LH	F	F	F
520	저녁	cənjək	2	LH	F	F	F
521	조금	cokɨm	2	LH	F	F	F
522	새벽	sepjək	2	LH	F	F	F
523	소금	sokɨm	2	LH	F	F	F
524	가을	kaɨl	2	LH	F	F	F
525	가방	kapaŋ	2	LH	F	F	F
526	마당	mataŋ	2	LH	F	F	F
527	기쁨	kippɨm	2	LH	F	F	F
528	무릎	murɨp	2	LH	F	F	F
529	구멍	kuməŋ	2	LH	F	F	F
530	스승	sɨsɨŋ	2	LH	F	F	F
531	부엌	puək	2	LH	F	F	F
532	대목	temok	2	LH	F	F	F
533	그늘	kɨnɨl	2	LH	F	F	F
534	지붕	cipuŋ	2	LH	F	F	F
535	구석	kusək	2	LH	F	F	F
536	바깥	pakkat	2	LH	F	F	F
537	보람	poram	2	LH	F	F	F
538	주먹	cumək	2	LH	F	F	F
539	여럿	jərət	2	LH	F	F	F
540	기둥	kituŋ	2	LH	F	F	F

541	나물	namul	2	LH	F	F	F
542	도둑	totuk	2	LH	F	F	F
543	뚜껑	ttukkəŋ	2	LH	F	F	F
544	가난	kanan	2	LH	F	F	F
545	가죽	kacuk	2	LH	F	F	F
546	계집	kjecip	2	LH	F	F	F
547	꾸중	kkucuŋ	2	LH	F	F	F
548	오줌	ocum	2	LH	F	F	F
549	사냥	sanjaŋ	2	LH	F	F	F
550	자갈	cakal	2	LH	F	F	F
551	개울	keul	2	LH	F	F	F
552	벼랑	pjəraŋ	2	LH	F	P	F
553	고을	koɨl	2	LH	F	F	F
554	배꼽	pekkop	2	LH	F	F	F
555	거름	kərɨm	2	LH	F	F	F
556	마중	macuŋ	2	LH	F	F	F
557	어음	əɨm	2	LH	F	P	F
558	이삭	isak	2	LH	F	F	F
559	뜨락	ttɨrak	2	LH	F	P	F
560	내음	neɨm	2	LH	F	F	F
561	아낙	anak	2	LH	F	F	F
562	지름	cirɨm	2	LH	F	F	F
563	도랑	toraŋ	2	LH	F	F	F
564	머슴	məsɨm	2	LH	F	F	F
565	어안	əan	2	LH	F	F	F
566	파랑	pharaŋ	2	LH	F	F	F
567	노을	noɨl	2	LH	F	F	F
568	허물	həmul	2	LH	F	F	F
569	허울	həul	2	LH	F	F	F
570	꺼풀	kkəphul	2	LH	F	F	F
571	노랑	noraŋ	2	LH	F	F	F
572	배짱	peccaŋ	2	LH	F	F	F
573	채찍	checcik	2	LH	F	F	D
574	메밀	memil	2	LH	F	F	F
575	노름	norɨm	2	LH	F	F	F
576	매듭	metɨp	2	LH	F	F	F
577	소쩍	soccək	2	LH	F	F	F
578	요강	jokaŋ	2	LH	F	F	F
579	투정	thucəŋ	2	LH	F	P	F

580	부럼	purəm	2	LH	F	F	F
581	으뜸	ɨttɨm	2	LH	F	F	F
582	주걱	cukək	2	LH	F	F	F
583	거죽	kəcuk	2	LH	F	F	F
584	두렁	turəŋ	2	LH	F	F	F
585	주검	cukəm	2	LH	F	F	F
586	개암	keam	2	LH	F	F	F
587	겨냥	kjənjaŋ	2	LH	F	F	D
588	버선	pəsən	2	LH	F	F	F
589	비탈	pithal	2	LH	F	F	F
590	소꿉	sokkup	2	LH	F	F	F
591	쪼끔	ccokkɨm	2	LH	F	F	F
592	헤엄	heəm	2	LH	F	F	F
593	고랑	koraŋ	2	LH	F	F	F
594	터울	thəul	2	LH	F	F	F
595	고명	komjəŋ	2	LH	F	F	F
596	수렁	surəŋ	2	LH	F	F	F
597	아양	ajaŋ	2	LH	F	F	F
598	우엉	uəŋ	2	LH	F	F	F
599	재갈	cekal	2	LH	F	F	F
600	호강	hokaŋ	2	LH	F	F	F
601	뜨물	ttɨmul	2	LH	F	F	F
602	마실	masil	2	LH	F	F	F
603	아범	apəm	2	LH	F	F	F
604	아욱	auk	2	LH	F	F	F
605	여물	jəmul	2	LH	F	F	F
606	여울	jəul	2	LH	F	F	F
607	이랑	iraŋ	2	LH	F	F	F
608	하양	hajaŋ	2	LH	F	F	F
609	가늠	kanɨm	2	LH	F	F	F
610	건사	kənsa	2	LH	F	F	F
611	고둥	kotuŋ	2	LH	F	F	F
612	구렁	kurəŋ	2	LH	F	F	F
613	기장	kicaŋ	2	LH	F	F	F
614	까탈	kkathal	2	LH	F	P	P
615	나방	napaŋ	2	LH	F	F	F
616	나염	najəm	2	LH	F	F	F
617	버들	pətɨl	2	LH	F	F	F
618	소댕	soteŋ	2	LH	F	F	?

619	어룽	əruŋ	2	LH	F	F	F
620	오름	orɨm	2	LH	F	F	F
621	자곱	cakop	2	LH	F	F	?
622	재첩	cechəp	2	LH	F	P	P
623	다음	taɨm	2	LH	P	P	P
624	모습	mosɨp	2	LH	P	P	P
625	이름	irɨm	2	LH	P	P	P
626	아들	atɨl	2	LH	P	P	P
627	오늘	onɨl	2	LH	P	P	P
628	하늘	hanɨl	2	LH	P	P	P
629	가슴	kasɨm	2	LH	P	P	P
630	느낌	nɨkkim	2	LH	P	P	P
631	여름	jərɨm	2	LH	P	P	P
632	겨울	kjəul	2	LH	P	P	P
633	이웃	iut	2	LH	P	P	P
634	까닭	kkatak	2	LH	P	P	P
635	바탕	pathaŋ	2	LH	P	P	P
636	며칠	mjəchil	2	LH	P	P	D
637	바닥	patak	2	LH	P	P	P
638	그릇	kirɨt	2	LH	P	P	P
639	마늘	manɨl	2	LH	P	P	P
640	기름	kirɨm	2	LH	P	P	P
641	노릇	norɨt	2	LH	P	P	P
642	거울	kəul	2	LH	P	P	P
643	마련	marjən	2	LH	P	P	P
644	버릇	pərɨt	2	LH	P	P	P
645	이틀	ithɨl	2	LH	P	P	P
646	바늘	panɨl	2	LH	P	P	P
647	이불	ipul	2	LH	P	P	P
648	스님	sɨnim	2	LH	P	P	P
649	조각	cokak	2	LH	P	P	P
650	사흘	sahɨl	2	LH	P	P	P
651	짜증	ccacɨŋ	2	LH	P	P	P
652	거짓	kəcit	2	LH	D	P	D
653	거품	kəphum	2	LH	P	P	P
654	구실	kusil	2	LH	P	P	P
655	화살	hwasal	2	LH	P	P	P
656	자랑	caraŋ	2	LH	P	P	P

657	바닥	patak	2	LH	P	P	P
658	토막	thomak	2	LH	P	P	P
659	자국	cakuk	2	LH	P	P	P
660	가닥	katak	2	LH	P	P	P
661	사슴	sasɨm	2	LH	P	P	P
662	씨름	ssirɨm	2	LH	P	P	P
663	보름	porɨm	2	LH	P	P	P
664	이슬	isɨl	2	LH	P	P	P
665	버섯	pəsət	2	LH	P	P	P
666	주름	curɨm	2	LH	P	P	P
667	씨앗	ssiat	2	LH	P	P	P
668	기슭	kisɨk	2	LH	P	P	F
669	하품	haphum	2	LH	P	P	P
670	기침	kichim	2	LH	P	P	P
671	기틀	kithɨl	2	LH	P	P	P
672	벼슬	pjəsɨl	2	LH	P	P	P
673	우물	umul	2	LH	P	P	P
674	자락	carak	2	LH	P	P	P
675	바둑	patuk	2	LH	P	P	P
676	저울	cəul	2	LH	P	P	P
677	사슬	sasɨl	2	LH	P	P	P
678	외상	wesaŋ	2	LH	P	P	P
679	기척	kichək	2	LH	P	P	P
680	호통	hothoŋ	2	LH	P	P	P
681	나흘	nahɨl	2	LH	P	P	P
682	다발	tapal	2	LH	P	P	P
683	미역	mijək	2	LH	P	P	P
684	비늘	pinɨl	2	LH	P	P	P
685	서슬	səsɨl	2	LH	P	P	P
686	어름	ərɨm	2	LH	P	P	P
687	고름 1(‘tie’)	korɨm	2	LH	P	P	P
688	티끌	thikkɨl	2	LH	P	P	D
689	구슬	kusɨl	2	LH	P	P	P
690	시름	sirɨm	2	LH	P	P	P
691	싸움	ssaum	2	LH	P	P	P
692	가락 (‘stick’)	karak	2	LH	P	P	P/F
693	그릇	kɨrɨt	2	LH	P	P	P
694	벼룩	pjəruk	2	LH	P	P	P

695	허탕	həthaŋ	2	LH	P	P	P
696	거북	kəpuk	2	LH	P	P	P
697	고름 2 (‘pus’)	korɨm	2	LH	P	D	D
698	지각	cikak	2	LH	P	P	P
699	꼬막	kkomak	2	LH	P	P	P
700	나락	narak	2	LH	P	P	P
701	수발	supal	2	LH	P	P	P
702	수선	susən	2	LH	P	P	P
703	고동	kotoŋ	2	LH	P	P	F
704	노끈	nokkɨn	2	LH	P	P	P
705	도막	tomak	2	LH	P	P	P
706	모냥	monjaŋ	2	LH	P	P	P
707	뽀록	pporok	2	LH	P	P	F
708	주책	cuchek	2	LH	P	P	D
709	고물(‘bean flour’)	komul	2	LH	P	P	D
710	고뿔	koppul	2	LH	P	P	P
711	더덕	tətək	2	LH	P	P	P
712	두릅	turɨp	2	LH	P	P	P
713	버짐	pəcim	2	LH	P	P	P
714	수술	susul	2	LH	P	P	P
715	오얏	ojat	2	LH	P	F	F
716	자반	capan	2	LH	P	P	D
717	자슥	casɨk	2	LH	P	P	P
718	쪼각	ccokak	2	LH	P	P	P
719	터럭	thərək	2	LH	P	P	?
720	튀각	thwikak	2	LH	P	P	P
721	튀밥	thwipap	2	LH	P	P	P
722	생각	seŋkak	2	HH	D	D	D
723	말씀	malssɨm	2	HH	D	D	D
724	걱정	kəkcəŋ	2	HH	D	D	D
725	한글	haŋkɨl	2	HH	D	D	D
726	한참	hancham	2	HH	D	D	D
727	신발	simpal	2	HH	D	D	D
728	골목	kolmok	2	HH	D	D	D
729	양념	jaŋnjəm	2	HH	D	D	D
730	장난	caŋnan	2	HH	D	D	D
731	임금	imkɨm	2	HH	D	D	D
732	말썽	malssəŋ	2	HH	D	D	D
733	엉망	əŋmaŋ	2	HH	D	D	D
734	물감	mulkam	2	HH	D	D	D

735	한창	hanchaŋ	2	HH	D	D	D
736	몰골	molkol	2	HH	D	D	D
737	속셈	soksem	2	HH	D	D	D
738	장승	caŋsɨŋ	2	HH	D	D	D
739	굴뚝	kulttuk	2	HH	D	D	D
740	딴판	ttamphan	2	HH	D	D	D
741	박살	paksal	2	HH	D	D	D
742	성냥	səŋnjaŋ	2	HH	D	D	D
743	동정	toŋcəŋ	2	HH	D	D	D
744	빗장	pitcaŋ	2	HH	D	D	D
745	헝겊	həŋkəp	2	HH	D	D	D
746	송장	soŋcaŋ	2	HH	D	D	D
747	익살	iksal	2	HH	D	D	D
748	법석	pəpsək	2	HH	D	D	D
749	송곳	soŋkot	2	HH	D	D	D
750	장님	caŋnim	2	HH	D	D	D
751	솔개	solke	2	HH	D	D	D
752	갑절	kapcəl	2	HH	D	D	D
753	곱절	kopcəl	2	HH	D	D	D
754	곱창	kopchaŋ	2	HH	D	D	D
755	장생	caŋseŋ	2	HH	D	D	D
756	젓갈	cətkal	2	HH	D	D	D
757	눈금	nuŋkɨm	2	HH	D	D	D
758	딴청	ttanchəŋ	2	HH	D	D	D
759	빙신	piŋsin	2	HH	D	D	D
760	장땡	caŋtteŋ	2	HH	D	D	D
761	간만	kamman	2	HH	D	D	D
762	골탕	kolthaŋ	2	HH	D	D	D
763	동냥	toŋnjaŋ	2	HH	D	D	D
764	안창	anchaŋ	2	HH	D	D	D
765	점잔	cəmcan	2	HH	D	D	D
766	칠갑	chilkap	2	HH	D	D	D
767	곱창	kopchaŋ	2	HH	D	D	D
768	길쌈	kilssam	2	HH	D	D	D
769	날품	nalphum	2	HH	D	D	P
770	목간	mokkan	2	HH	D	D	D
771	억척	əkchək	2	HH	D	D	D
772	엽삼	jəpsam	2	HH	D	D	D
773	젬병	cempjəŋ	2	HH	D	D	P

774	진동	cintoŋ	2	HH	D	D	D
775	얼굴	əlkul	2	HH	F	F	F
776	동생	toŋseŋ	2	HH	F	F	F
777	잘못	calmot	2	HH	F	F	F
778	입술	ipsul	2	HH	F	F	F
779	달걀	talkjal	2	HH	F	F	F
780	짐승	cimsɨŋ	2	HH	F	F	F
781	잠깐	camkkan	2	HH	F	F	F
782	당근	taŋkɨn	2	HH	F	F	F
783	살림	sallim	2	HH	F	F	F
784	언덕	əntək	2	HH	F	F	F
785	영문	jəŋmun	2	HH	F	F	F
786	받침	patchim	2	HH	F	F	F
787	검정	kəmcəŋ	2	HH	F	F	F
788	덩굴	təŋkul	2	HH	F	F	F
789	근심	kɨnsim	2	HH	F	F	F
790	신명	simmjəŋ	2	HH	F	F	F
791	연장	jəncaŋ	2	HH	F	F	F
792	빨강	ppalkaŋ	2	HH	F	F	F
793	장단	caŋtan	2	HH	F	D	F
794	말뚝	malttuk	2	HH	F	F	F
795	건성	kənsəŋ	2	HH	F	F	F
796	징검	ciŋkəm	2	HH	F	P	F
797	헛간	həkkan	2	HH	F	F	F
798	얼룩	əlluk	2	HH	F	F	F
799	옹달	oŋtal	2	HH	F	P	F
800	천둥	chəntuŋ	2	HH	F	F	F
801	앙금	aŋkɨm	2	HH	F	D	F
802	흥정	hɨŋcəŋ	2	HH	F	F	F
803	선반	səmpan	2	HH	F	P	F
804	잔등	cantɨŋ	2	HH	F	F	F
805	넝쿨	nəŋkhul	2	HH	F	F	F
806	안심	ansim	2	HH	F	F	F
807	응달	ɨŋtal	2	HH	F	P	P
808	능청	nɨŋchəŋ	2	HH	F	F	F
809	멍석	məŋsək	2	HH	F	F	F
810	골짝	kolccak	2	HH	F	F	F
811	덤불	təmpul	2	HH	F	F	F
812	방정	paŋcəŋ	2	HH	F	F	F

813	뽕짝	ppoŋccak	2	HH	F	F	F
814	양푼	jaŋphun	2	HH	F	F	F
815	건방	kəmpaŋ	2	HH	F	F	F
816	궁창	kuŋchaŋ	2	HH	F	F	F
817	깜장	kkamcaŋ	2	HH	F	F	F
818	남방	nampaŋ	2	HH	F	F	F
819	능금	nɨŋkɨm	2	HH	F	F	F
820	멍울	məŋul	2	HH	F	F	F
821	멍청	məŋchəŋ	2	HH	F	F	F
822	방천	paŋchən	2	HH	F	F	F
823	봉창	poŋchaŋ	2	HH	F	F	F
824	삽짝	sapccak	2	HH	F	F	F
825	엄장	əmcaŋ	2	HH	F	F	F
826	손님	sonnim	2	HH	P	P	P
827	목숨	moksum	2	HH	P	P	P
828	한숨	hansum	2	HH	P	P	P
829	껍질	kkəpcil	2	HH	P	P	P
830	방울	paŋul	2	HH	P	P	P
831	반죽	pancuk	2	HH	P	P	P
832	멱살	mjəksal	2	HH	P	P	P
833	단골	tankol	2	HH	P	P	P
834	숭늉	suŋnjuŋ	2	HH	P	P	P
835	범벅	pəmpək	2	HH	P	P	P
836	쪽박	ccokpak	2	HH	P	P	F
837	쑥갓	ssukkat	2	HH	P	P	P
838	조기	coki	2	HH	P	P	P
839	늑장	nɨkcaŋ	2	HH	P	P	P
840	방죽	paŋcuk	2	HH	P	P	?
841	엄살	əmsal	2	HH	P	P	P
842	콩팥	khoŋphat	2	HH	P	P	P
843	핀잔	phincan	2	HH	P	P	P
844	빗금	pikkɨm	2	HH	P	P	P
845	안달	antal	2	HH	P	P	P
846	앙탈	aŋthal	2	HH	P	P	P
847	전골	cəŋkol	2	HH	P	P	P
848	강정	kaŋcəŋ	2	HH	P	P	P
849	넉살	nəksal	2	HH	P	P	P
850	둔덕	tuntək	2	HH	P	P	D
851	응석	ɨŋsək	2	HH	P	P	P

852	볼품	polphum	2	HH	P	P	P
853	장만	caŋman	2	HH	P	P	P
854	동강	toŋkaŋ	2	HH	P	P	P
855	먹통	məkthoŋ	2	HH	P	P	D
856	불알	pural	2	LH	P	P	P
857	진창	cinchaŋ	2	HH	P	P	D
858	둔턱	tunthək	2	HH	P	P	D
859	등걸	tɨŋkəl	2	HH	P	P	P
860	등골	tɨŋkol	2	HH	P	P	P
861	삿갓	satkat	2	HH	P	P	P
862	영각	jəŋkak	2	HH	P	P	?
863	응석	ɨŋsək	2	HH	P	P	P
864	장목	caŋmok	2	HH	P	P	P
865	풍장	phuŋcaŋ	2	HH	P	P	P
866	이야기	ijaki	3	LLL	D	D	D
867	마누라	manura	3	LLL	D	D	D
868	고구마	kokuma	3	LLL	D	D	D
869	다래끼	tarekki	3	LLL	P	P	P
870	사투리	sathuri	3	LLL	D	D	D
871	보조개	pocoke	3	LLL	D	D	D
872	거머리	kəməri	3	LLL	D	D	D
873	무지개	mucike	3	LLL	D	D	D
874	구더기	kutəki	3	LLL	D	D	D
875	기지개	kicike	3	LLL	D	D	D
876	오라비	orapi	3	LLL	D	D	D
877	새내기	seneki	3	LLL	D	D	D
878	새치기	sechiki	3	LLL	D	D	I
879	사마귀	samakwi	3	LLL	D	D	D
880	꼬다리	kkotari	3	LLL	D	D	P
881	꽈배기	kkwapeki	3	LLL	D	D	D
882	도떼기	totteki	3	LLL	I	D	I
883	이무기	imuki	3	LLL	D	D	D
884	이바구	ipaku	3	LLL	D	D	D
885	자치기	cachiki	3	LLL	D	D	D
886	키다리	khitari	3	LLL	D	D	D
887	해파리	hephari	3	LLL	D	D	I
888	사다리	satari	3	LLL	F	F	F
889	너스레	nəsɨre	3	LLL	F	F	F
890	허드레	hətɨre	3	LLL	F	F	F
891	가시나	kasina	3	LLL	F	F	F

892	며느리	mjənɨri	3	LLL	I	I	I
893	쏘시개	ssosike	3	LLL	I	P	P
894	아지매	acime	3	LLL	I	I	I
895	가마니	kamani	3	LLL	I	P	I
896	바리때	paritte	3	LLL	I/P	I	I
897	다래끼	tarekki	3	LLL	P	P	P
898	어머니	əməni	3	LLL	P	P	P
899	아버지	apəci	3	LLL	P	P	P
900	아저씨	acəssi	3	LLL	P	P	P
901	쓰레기	ssɨreki	3	LLL	P	P	P
902	나머지	naməci	3	LLL	P	P	P
903	스스로	sɨsɨro	3	LLL	P	P	P
904	주머니	cuməni	3	LLL	P	P	P
905	아가씨	akassi	3	LLL	P	P	P
906	저고리	cəkori	3	LLL	P	P	P
907	마무리	mamuri	3	LLL	P	P	P
908	제자리	cecari	3	LLL	P	P	P
909	테두리	theturi	3	LLL	P	P	P
910	개구리	kekuri	3	LLL	P	P	P
911	찌꺼기	ccikkəki	3	LLL	P	P	P
912	기저귀	kicəki	3	LLL	P	P	P
913	바구니	pakuni	3	LLL	P	P	P
914	사나이	sanai	3	LLL	P	P	P
915	무더기	mutəki	3	LLL	P	P	P
916	나그네	nakɨne	3	LLL	P	P	P
917	도깨비	tokkepi	3	LLL	P	P	P
918	나들이	natɨri	3	LLL	P	P	P
919	오누이	onui	3	LLL	P	P	P
920	소나기	sonaki	3	LLL	P	P	P
921	싸구려	ssakurjə	3	LLL	P	P	F
922	개나리	kenari	3	LLL	P	P	P
923	모서리	mosəri	3	LLL	P	P	P
924	바가지	pakaci	3	LLL	P	P	P
925	수제비	sucepi	3	LLL	P	P	P
926	대가리	tekari	3	LLL	P	P	P
927	구두쇠	kutuswe	3	LLL	P	P	P
928	도라지	toraci	3	LLL	P	P	P
929	꼬투리	kkothuri	3	LLL	P	P	P
930	소쿠리	sokhuri	3	LLL	P	P	P
931	재채기	cecheki	3	LLL	P	I	I
932	고사리	kosari	3	LLL	P	P	P

933	코끼리	khokkiri	3	LLL	P	P	P
934	꾸러미	kkurəmi	3	LLL	P	I	P
935	미나리	minari	3	LLL	P	P	P
936	지우개	ciuke	3	LLL	P	P	P
937	다시마	tasima	3	LLL	P	P	P
938	도토리	tothori	3	LLL	P	P	P
939	메뚜기	mettuki	3	LLL	P	P	P
940	뻐꾸기	ppəkkuki	3	LLL	P	P	P
941	이파리	iphari	3	LLL	P	P	P
942	시치미	sichimi	3	LLL	P	P	P
943	귀고리	kwikori	3	LLL	P	P	P
944	기러기	kirəki	3	LLL	P	P	P
945	까투리	kkathuri	3	LLL	P	P	P
946	느타리	nɨthari	3	LLL	P	P	P
947	수세미	susemi	3	LLL	P	P	P
948	시래기	sireki	3	LLL	P	P	P
949	싸가지	ssakaci	3	LLL	P	P	P
950	누더기	nutəki	3	LLL	P	P	P
951	두더지	tutəci	3	LLL	P	P	P
952	메아리	meari	3	LLL	P	P	P
953	서까래	səkkare	3	LLL	P	P	P
954	피라미	phirami	3	LLL	P	P	P
955	회초리	hwechori	3	LLL	P	P	P
956	꾀꼬리	kkwekkori	3	LLL	P	P	P
957	어버이	əpəi	3	LLL	P	P	P
958	허깨비	həkkəpi	3	LLL	P	P	P
959	까마귀	kkamakwi	3	LLL	P	P	P
960	너구리	nəkuri	3	LLL	P	P	P
961	모가지	mokaci	3	LLL	P	P	P
962	배재기	peceki	3	LLL	P	P	?
963	자투리	cathuri	3	LLL	P	P	P
964	회오리	hweori	3	LLL	P	P	P
965	그저께	kɨcəkke	3	LLL	P	P	P
966	꾸러기	kkurəki	3	LLL	P	I	P
967	다리미	tarimi	3	LLL	P	P	P
968	두꺼비	tukkəpi	3	LLL	P	P	P
969	오마니	omani	3	LLL	P	P	P
970	쪼가리	ccokari	3	LLL	P	P	P
971	포대기	photeki	3	LLL	P	P	P
972	노리개	norike	3	LLL	P	P	P
973	또아리	ttoari	3	LLL	P	P	P

974	소구리	sokuri	3	LLL	P	P	P
975	아가미	akami	3	LLL	P	P	P
976	우거지	ukəci	3	LLL	P	P	P
977	우스개	usɨke	3	LLL	P	P	P
978	지아비	ciapi	3	LLL	P	P	P
979	가리비	karipi	3	LLL	P	P	P
980	도가니	tokani	3	LLL	P	P	P
981	뜨내기	ttɨneki	3	LLL	P	P	P
982	모내기	moneki	3	LLL	P	P	P
983	미닫이	mitaci	3	LLL	P	P	P
984	바구미	pakumi	3	LLL	P	P	P
985	부꾸미	pukkumi	3	LLL	P	P	P
986	가두리	katuri	3	LLL	P	P	P
987	노가리	nokari	3	LLL	P	P	P
988	도라이	torai	3	LLL	P	P	P
989	도리깨	torikke	3	LLL	P	P	P
990	떠돌이	ttətori	3	LLL	P	P	P
991	마타리	mathari	3	LLL	P	P	P
992	머저리	məcəri	3	LLL	P	P	P
993	무서리	musəri	3	LLL	P	P	P
994	미투리	mithuri	3	LLL	P	P	P
995	버러지	pərəci	3	LLL	P	P	P
996	소보루	soporu	3	LLL	P	P	P
997	소쩍이	soccəki	3	LLL	P	P	P
998	싸라기	ssaraki	3	LLL	P	P	P
999	에누리	enuri	3	LLL	P	P	P
1000	오뚝이	ottuki	3	LLL	P	P	P
1001	오소리	osori	3	LLL	P	P	P
1002	오자미	ocami	3	LLL	P	P	P
1003	이바지	ipaci	3	LLL	P	P	P
1004	가오리	kaori	3	LLL	P	P	I
1005	가자미	kacami	3	LLL	P	P	P
1006	구루마	kuruma	3	LLL	P	P	P
1007	꼬라지	kkoraci	3	LLL	P	P	P
1008	노다지	notaci	3	LLL	P	P	P
1009	도요새	tojose	3	LLL	P	P	P
1010	도지개	tocike	3	LLL	P	P	P
1011	두루미	turumi	3	LLL	P	P	P
1012	따개비	ttakepi	3	LLL	P	P	P
1013	마고자	makoca	3	LLL	P	P	P
1014	마구리	makuri	3	LLL	P	P	P

1015	매가리	mekari	3	LLL	P	P	P
1016	바지게	pacike	3	LLL	P	P	P
1017	배지기	peciki	3	LLL	P	P	P
1018	비짜루	piccaru	3	LLL	P	P	P
1019	싸리비	ssaripi	3	LLL	P	P	P
1020	아니리	aniri	3	LLL	P	P	P
1021	아재비	acepi	3	LLL	P	P	P
1022	오라기	oraki	3	LLL	P	P	P
1023	우수리	usuri	3	LLL	P	P	P
1024	자배기	capeki	3	LLL	P	P	P
1025	조가비	cokapi	3	LLL	P	P	P
1026	지게미	cikemi	3	LLL	P	P	P
1027	지어미	ciəmi	3	LLL	P	P	P
1028	쭈쭈바	ccuccupa	3	LLL	P	P	I
1029	허재비	həcepi	3	LLL	P	P	D
1030	휘모리	hwimori	3	LLL	P	P	P
1031	실마리	silmari	3	HLL	D	D	P
1032	한가위	haŋkawi	3	HLL	D	D	D
1033	올가미	olkami	3	HLL	D	P	P
1034	할아비	harapi	3	LLL	D	D	D
1035	얼치기	əlchiki	3	HLL	D	D	D
1036	박고지	pakkoci	3	HLL	D	D	D
1037	딴따라	ttanttara	3	HLL	D	D	D
1038	물푸레	mulphure	3	HLL	D	P	P
1039	뺑소니	ppeŋsoni	3	HLL	D	D	P
1040	각다귀	kaktakwi	3	HLL	D	P	D
1041	동치미	toŋchimi	3	HLL	D	D	D
1042	딱따기	ttakttaki	3	HLL	D	P	P
1043	멱서리	mjəksəri	3	HLL	D	P	P
1044	문디이	muntii	3	HLL	D	D	D
1045	심마니	simmani	3	HLL	D	D	D
1046	품앗이	phumasi	3	HLL	D	P	P
1047	잡도리	captori	3	HLL	P	I	P
1048	날치기	nalchiki	3	HLL	I	I	I
1049	할머니	halməni	3	HLL	P	P	D
1050	덩어리	təŋəri	3	HLL	P	P	P
1051	꼭대기	kkokteki	3	HLL	P	P	P
1052	동아리	toŋari	3	HLL	P	P	P
1053	골짜기	kolccaki	3	HLL	P	P	P
1054	잠자리	camcari	3	HLL	P	P	P
1055	울타리	ulthari	3	HLL	P	P	P

1056	옥수수	oksusu	3	HLL	P	P	P
1057	줄거리	culkəri	3	HLL	P	P	P
1058	설거지	səlkəci	3	HLL	P	P	P
1059	봉우리	poŋuri	3	HLL	P	P	P
1060	엉터리	əŋthəri	3	HLL	P	P	P
1061	잎사귀	ipsakwi	3	HLL	P	P	P
1062	송아지	soŋaci	3	HLL	P	P	P
1063	언저리	əncəri	3	HLL	P	P	P
1064	광주리	kwaŋcuri	3	HLL	P	P	P
1065	병아리	pjəŋari	3	HLL	P	P	P
1066	껍데기	kkəpteki	3	HLL	P	P	P
1067	종아리	coŋari	3	HLL	P	P	P
1068	막바지	makpaci	3	HLL	P	P	P
1069	빗자루	pitcaru	3	HLL	P	P	P
1070	꽁무니	kkoŋmuni	3	HLL	P	P	P
1071	막대기	makteki	3	HLL	P	P	P
1072	장도리	caŋtori	3	HLL	P	P	P
1073	벙어리	pəŋəri	3	HLL	P	P	P
1074	강아지	kaŋaci	3	HLL	P	P	P
1075	작대기	cakteki	3	HLL	P	P	P
1076	뚝배기	ttupeki	3	HLL	P	P	P
1077	방아쇠	paŋaswe	3	HLL	P	P	P
1078	응어리	ɨŋəri	3	HLL	P	P	P
1079	돋보기	totpoki	3	HLL	P	P	P
1080	망아지	maŋaci	3	HLL	P	P	P
1081	단무지	tammuci	3	HLL	P	P	P
1082	생채기	seŋcheki	3	HLL	P	P	P
1083	송사리	soŋsari	3	HLL	P	P	P
1084	홍두깨	hoŋtukke	3	HLL	P	P	P
1085	달구지	talkuci	3	HLL	P	P	P
1086	진드기	cintɨki	3	HLL	P	P	P
1087	진저리	cincəri	3	HLL	P	P	P
1088	날라리	nallari	3	HLL	P	P	P
1089	장아찌	caŋacci	3	HLL	P	P	P
1090	건더기	kəntəki	3	HLL	P	P	P
1091	들러리	tɨlləri	3	HLL	P	P	P
1092	번데기	pənteki	3	HLL	P	P	P
1093	삼태기	samtheki	3	HLL	P	P	P
1094	장다리	caŋtari	3	HLL	P	P	P

1095	족두리	cokturi	3	HLL	P	P	P
1096	족제비	cokcepi	3	HLL	P	P	P
1097	쫀드기	ccontɨki	3	HLL	P	P	P
1098	갈고리	kalkori	3	HLL	P	P	P
1099	곱빼기	kopppeki	3	HLL	P	P	P
1100	깡다구	kkaŋtaku	3	HLL	P	P	P
1101	낟가리	natkari	3	HLL	P	P	P
1102	덧게비	tətkepi	3	HLL	P	P	P
1103	둥우리	tuŋuri	3	HLL	P	P	P
1104	똘마니	ttolmani	3	HLL	P	P	P
1105	망나니	maŋnani	3	HLL	P	P	P
1106	멱아지	mjəkaci	3	HLL	P	P	P
1107	봉다리	poŋtari	3	HLL	P	P	P
1108	봉오리	poŋori	3	HLL	P	P	P
1109	살무사	salmusa	3	HLL	P	P	P
1110	상수리	saŋsuri	3	HLL	P	D	P
1111	송두리	soŋturi	3	HLL	P	D	P
1112	씀바귀	ssɨmpakwi	3	HLL	P	P	P
1113	얼루기	əlluki	3	HLL	P	P	P
1114	중모리	cuŋmori	3	HLL	P	P	P
1115	짝짝이	ccakccaki	3	HLL	P	P	P
1116	쫄따구	ccolttaku	3	HLL	P	P	P
1117	퉁구리	thuŋkuri	3	HLL	P	P	?
1118	그림자	kɨrimca	3	LHL	D	D	D
1119	뒤꿈치	twikkumchi	3	LHL	D	D	D
1120	허벅지	həpəkci	3	LHL	D	D	D
1121	지렁이	cirəŋi	3	LHL	D	D	D
1122	구렁이	kurəŋi	3	LHL	D	D	D
1123	푸성귀	phusəŋkwi	3	LHL	D	D	P
1124	개망초	kemaŋcho	3	LHL	D	D	P
1125	가운데	kaunte	3	LHL	P	P	P
1126	어린이	ərini	3	LLL	P	P	P
1127	고양이	kojaŋi	3	LHL	P	P	P
1128	오징어	ociŋə	3	LHL	P	P	P
1129	아줌마	acumma	3	LHL	P	P	P
1130	모퉁이	mothuŋi	3	LHL	P	P	P
1131	누룽지	nuruŋci	3	LHL	P	P	P
1132	아낙네	anakne	3	LHL	P	P	P
1133	귀퉁이	kwithuŋi	3	LHL	P	P	P

1134	시금치	sikɨmchi	3	LHL	P	P	P
1135	구덩이	kutəŋi	3	LHL	P	P	P
1136	비둘기	pitulki	3	LHL	P	P	P
1137	아궁이	akuŋi	3	LHL	P	P	P
1138	버들치	pətɨlchi	3	LHL	P	P	P
1139	자물쇠	camulswe	3	LHL	P	P	P
1140	거문고	kəmunko	3	LHL	P	P	P
1141	다람쥐	taramcwi	3	LHL	P	P	P
1142	주둥이	cutuŋi	3	LHL	P	P	P
1143	지팡이	ciphaŋi	3	LHL	P	P	P
1144	가랑이	karaŋi	3	LHL	P	P	P
1145	가물치	kamulchi	3	LHL	P	P	I
1146	어금니	əkɨmni	3	LHL	P	P	P
1147	거북이	kəpuki	3	LLL	P	P	P
1148	누린내	nurinne	3	LHL	P	P	P
1149	주근깨	cukɨnkke	3	LHL	P	P	P
1150	추임새	chuimse	3	LHL	P	P	P
1151	고등어	kotɨŋə	3	LHL	P	P	P
1152	꼬맹이	kkomeŋi	3	LHL	P	P	P
1153	오랑캐	oraŋkhe	3	LHL	P	P	P
1154	외톨이	wetholi	3	LHL	P	P	P
1155	자전거	cacənkə	3	LHL	P	P	P
1156	꺼벙이	kkəpəŋi	3	LHL	P	P	P
1157	꼬챙이	kkocheŋi	3	LHL	P	P	P
1158	노랑이	noraŋi	3	LHL	P	P	P
1159	도롱이	toroŋi	3	LHL	P	P	P
1160	미장이	micaŋi	3	LHL	P	P	P
1161	바랑이	paraŋi	3	LHL	P	P	P
1162	바랭이	pareŋi	3	LHL	P	P	P
1163	벼멸구	pjəmjəlku	3	LHL	P	P	P
1164	오쟁이	oceŋi	3	LHL	P	P	P
1165	고쟁이	koceŋi	3	LHL	P	P	P
1166	꼬랑지	kkoraŋci	3	LHL	P	P	P
1167	다랑이	taraŋi	3	LHL	P	P	P
1168	바탱이	patheŋi	3	LHL	P	P	P
1169	부엉이	puəŋi	3	LHL	P	P	P
1170	부쟁이	puceŋi	3	LHL	P	P	?

1171	애송이	esoŋi	3	LHL	P	D	D
1172	우듬지	utɨmci	3	LHL	P	P	?
1173	지짐이	cicimi	3	LLL	P	P	P
1174	마나님	mananim	3	LLH	D	D	D
1175	하소연	hasojən	3	LLH	D	D	D
1176	어르신	ərɨsin	3	LLH	D	D	D
1177	여드름	jətɨrɨm	3	LLH	D	D	D
1178	오지랖	ocirap	3	LLH	D	F	F
1179	고드름	kotɨrɨm	3	LLH	D	D	D
1180	마지막	macimak	3	LLH	F	F	F
1181	도시락	tocirak	3	LLH	F	F	F
1182	조바심	copasim	3	LLH	F	F	F
1183	무르팍	murɨphak	3	LLH	F	F	P
1184	거드름	kətɨrɨm	3	LLH	F	D	F
1185	부뚜막	puttumak	3	LLH	F	F	F
1186	어리광	ərikwaŋ	3	LLH	F	F	F
1187	부스럼	pusɨrəm	3	LLH	F	F	F
1188	꼬부랑	kkopuraŋ	3	LLH	F	F	F
1189	도루묵	torumuk	3	LLH	F	F	F
1190	바지락	pacirak	3	LLH	F	F	F
1191	수수깡	susukkaŋ	3	LLH	F	I	F
1192	하느님	hanɨnim	3	LLH	I	I	I
1193	하나님	hananim	3	LLH	I	I	I
1194	요즈음	jocɨɨm	3	LLH	P	P	P
1195	휘파람	hwipharam	3	LLH	P	D	D
1196	아리랑	ariraŋ	3	LLH	P	P	P
1197	뜨개질	ttɨkecil	3	LLH	P	P	P
1198	끄나풀	kkɨnaphul	3	LLH	P	P	P
1199	엉덩이	əŋtəŋi	3	HHL	D	D	D
1200	원숭이	wənsuŋi	3	HHL	D	D	D
1201	곰팡이	komphaŋi	3	HHL	D	P	D
1202	궁둥이	kuŋtuŋi	3	HHL	D	D	D
1203	웅덩이	uŋtəŋi	3	HHL	D	I	I
1204	문둥이	muntuŋi	3	HHL	D	D	D
1205	널빤지	nəlppanci	3	HHL	D	P	D
1206	곡괭이	kokkweŋi	3	HHL	D	P	D
1207	뚱딴지	ttuŋttanci	3	HHL	D	D	I
1208	곰방대	kompaŋte	3	HHL	D	P	D

1209	얼간이	əlkani	3	HLL	D	D	D
1210	굼벵이	kumpeŋi	3	HHL	D	D	D
1211	난쟁이	nanceŋi	3	HHL	D	D	D
1212	장딴지	caŋttanci	3	HHL	D	D	D
1213	곰장어	komcaŋə	3	HHL	D	P	P
1214	난장이	nancaŋi	3	HHL	D	D	D
1215	맹꽁이	meŋkkoŋi	3	HHL	D	D	D
1216	살쾡이	salkhweŋi	3	HHL	D	P	P
1217	접영초	cəpjəŋcho	3	LHL	D	D	D
1218	좀팽이	compheŋi	3	HHL	D	D	P
1219	진달래	cintalle	3	HHL	F	F	F
1220	복숭아	poksuŋa	3	HHL	F	F	F
1221	민들레	mintɨlle	3	HHL	F	F	F
1222	둥굴레	tuŋkulle	3	HHL	F	F	F
1223	신들메	sintɨlme	3	HHL	F	P	P
1224	막걸리	makkəlli	3	HHL	P	P	P
1225	돌멩이	tolmeŋi	3	HHL	P	P	P
1226	몽둥이	moŋtuŋi	3	HHL	P	P	P
1227	알갱이	alkeŋi	3	HHL	P	P	P
1228	실랑이	sillaŋi	3	HHL	P	P	P
1229	알맹이	almeŋi	3	HHL	P	P	P
1230	질경이	cilkjəŋi	3	HHL	P	P	P
1231	방망이	paŋmaŋi	3	HHL	P	P	P
1232	청경채	chəŋkjəŋche	3	HHL	P	P	P
1233	맞장구	matcaŋku	3	HHL	P	P	P
1234	달팽이	talpheŋi	3	HHL	P	P	P
1235	올챙이	olcheŋi	3	HHL	P	P	P
1236	등성이	tɨsəŋi	3	HHL	P	P	P
1237	삭정이	sakcəŋi	3	HHL	P	P	P
1238	인절미	incəlmi	3	HHL	P	P	P
1239	종달새	coŋtalse	3	HHL	P	P	P
1240	강냉이	kaŋneŋi	3	HHL	P	P	P
1241	굴렁쇠	kulləŋswe	3	HHL	P	P	P
1242	끈끈이	kkɨnkkɨni	3	HLL	P	P	P
1243	양송이	jaŋsoŋi	3	HHL	P	P	P
1244	잠방이	campaŋi	3	HHL	P	P	P

1245	정강이	cəŋkaŋi	3	HHL	P	P	P
1246	골뱅이	kolpeŋi	3	HHL	P	P	P
1247	남생이	namseŋi	3	HHL	P	P	P
1248	담쟁이	tamceŋi	3	HHL	P	P	P
1249	빨갱이	ppalkeŋi	3	HHL	P	P	P
1250	승냥이	sɨŋnjaŋi	3	HHL	P	P	P
1251	족집게	ccokcipke	3	HHL	P	P	P
1252	풍뎅이	phuŋteŋi	3	HHL	P	P	P
1253	딸딸이	ttalttari	3	HLL	P	P	P
1254	뭉텅이	muŋthəŋi	3	HHL	P	P	P
1255	밴댕이	penteŋi	3	HHL	P	P	P
1256	업둥이	əptuŋi	3	HHL	P	P	P
1257	엉겅퀴	əŋkəŋkhwi	3	HHL	P	P	P
1258	쭉정이	ccukcəŋi	3	HHL	P	P	P
1259	저돌적	cətolcək	3	LHH	D	I	D
1260	외양간	wejaŋkan	3	LHH	F	F	F
1261	으름장	ɨrɨmcaŋ	3	LHH	F	F	F
1262	호들갑	hotɨlkap	3	LHH	F	F	F
1263	뒤웅박	twiuŋpak	3	LHH	F	F	P
1264	구렁텅	kurəŋthəŋ	3	LHH	F	F	F
1265	도롱뇽	torjoŋnjoŋ	3	LHH	F	F	F
1266	도련님	torjənnim	3	LHH	P	P	P
1267	두반장	tupancaŋ	3	LHH	P	P	P
1268	심부름	simpurɨm	3	HLH	D	D	D
1269	함지박	hamcipak	3	HLH	D	F	D
1270	왈가닥	walkatak	3	HLH	F	F	F
1271	엉덩짝	əŋtəŋccak	3	HHH	D	D	D
1272	방앗간	paŋatkan	3	HHH	F	F	F
1273	딸꾹질	ttalkukcil	3	HHH	P	P	P
1274	진딧물	cintitmul	3	HHH	P	P	P
1275	군것질	kunkətcil	3	HHH	P	P	P

*In the transcriptions, *ə* stands for the mid unrounded vowel, *ɨ* for the high central/back unrounded vowel, *c* for an alveolar affricate, geminate consonant (CC: eg., *tt*) for a tense consonant, and C*h* for an aspirated consonant.
** P stands for penultimate accent; D for double accent; F for final accent; I for initial accent; and ? for a response, "don't know".

Chapter 5 The Role of Analogy: Word Similarity Effects

5.1 Introduction

Chapter 3 demonstrated that syllable structure plays a role in NKK accent assignment in novel words. In Chapter 4, I argued that the syllable structure-sensitive patterns are actually a reflection of lexical statistical patterns and that NKK speakers internalize the lexical statistical patterning. Furthermore, the regular patterning appeared even more strongly in novel words than in the lexicon.

This chapter investigates whether another factor, analogy to phonetically similar words, crucially influences accent placement in novel words. Several different studies have reported that analogy to similar words plays a role in the patterning of novel words (e.g. Baker and Smith 1976; Guion *et al.* 2003; Eddington 2000; Face 2004). Baker and Smith (1976) investigated English stress patterns in nonwords, testing the sometimes conflicting patterns predicted by SPE stress rules vs. analogy to real words. They found that in many cases, stress was assigned based on the stress of similar words, even when the resulting stress patterns were in conflict with the general rules. Guion *et al.* (2003) also showed that the location of stress on phonologically similar real words influenced stress placement in English nonwords, although they claimed that analogy is just one of the factors affecting stress patterns, others including syllable structure and lexical class. Investigating the factors determining the perception of stress, Face (2004) performed a study in which synthesized productions of Spanish nonwords were controlled so that all syllables were equal in acoustic prominence (F0, duration, and intensity). Face examined several different factors including word similarity, syllable weight, and morphological category, and found that stress in phonetically similar words was one of the factors that influenced stress perception. On this view, language acquisition involves storing and categorizing linguistic experience, rather than generalizing from linguistic input to abstract rules or constraints, and novel words should be processed via analogy to existing words (Bybee 1985, 2001; Daelemans, Gillis, and Durieux 1994; Skousen 1989).

Because NKK lexical accent patterns are not predictable and therefore must be stored as part of a word's lexical entry, it seems reasonable to expect that accent patterns of existing similar words would play a major role in determining of the accent position that speakers assign in novel words. Therefore, this chapter presents an experimental study to test the *Word Similarity Effect Hypothesis*, which predicts that in assigning accent to novel words, native speakers will be influenced by the accent patterns of phonetically similar existing words. The results do not support this hypothesis: NKK speakers do not necessarily follow the accent patterns of analogous existing words when accenting novel words. Rather, they tend to accent novel words in accordance with syllable structure, even when an analogous existing word with the same syllable structure has a different accent pattern. The results clearly demonstrate that the effects of syllable structure outweigh the effects of word similarity in NKK accent patterns in novel words.

Section 5.2 of this chapter presents an experimental study of NKK accentuation in novel words that have phonetically similar counterparts. Section 5.3 discusses the findings and compares accent patterns in novel words which have obvious existing counterparts vs. novel words which do not (shown in Chapter 3). Section 5.4 is the conclusion.

5.2 Experiment 2

5.2.1 Participants and Procedure

The participants were the same as those in the first experiment presented in Chapter 3: 30 NKK speakers, ranging in age from 29 to 54 years (mean 39), 15 male and 15 female. The current experiment was performed right after the first experiment, following the same procedure as in Experiment 1.

5.2.2 Materials

The major difference between the experiment reported on in this chapter and that reported on in Chapter 3 is that the experiment in this chapter used novel words which were phonetically similar to existing words, whereas the earlier experiment used novel words which did not have obvious existing counterparts. In this experiment, the stimulus set consisted of novel words, 24 bisyllabic words and 24 trisyllabic words. The novel words were identical to existing words except for one consonant in the final syllable.[16] In most cases, the corresponding consonants in the existing vs. the novel word differed in only one place feature: for example, novel word *kɨri**ŋ*** vs. real word *kɨ́rí**m*** 'picture'; novel word *muci**d**e* vs. real word *múcí**g**e* 'rainbow'. Liquid sounds were replaced by nasal sounds at a different place of articulation (e.g. novel word *tango**ŋ*** vs. real word *tángo**l*** 'customer'). Following findings that the onset of a word activates a set of lexical candidates sharing that onset (e.g. Allopenna *et al*. 1998; Marslen-Wilson 1987), I assumed that the novel stimuli would activate neighboring lexical items.[17]

To determine whether analogy operates independently of syllable structure, each set of novel words corresponding to a particular syllable structure (e.g. CVCV) was balanced to ensure that each accent type was equally represented in the existing words that corresponded to the novel forms. The novel forms were separated into four different word types: for bisyllabic forms, LL, HL, LH, HH, and for trisyllabic forms, LLL, HLL, LLH, HHL (L=light syllable; H=heavy syllable). Six items were included for each word type: 2 corresponding to analogical real words with double accent, 2 corresponding to analogical real words with penultimate accent, and 2 corresponding to analogical real words with final accent. Due to the absence of final accent in existing HLL words, HLL novel forms targeting analogical final accent were not included. The novel words that were used in the experiment and the analogical real counterparts are presented in (30).

[16] One item *paŋraŋi* was created by changing a consonant in the penultimate syllable of a word *paŋ.maŋ.i.* 'cudgel' because the existing word did not contain a consonant in the final syllable (*ŋ* is not a possible onset so *ŋ* in the second syllable would not be resyllabified as onset of the following syllable).

[17] Additional analogous existing counterparts might be available for the novel forms used in Experiment 2. The possibility of competing analogical accent patterns is discussed in Section 5.4. Frequency of neighbors was not considered in this study because most of the target neighboring items were considered as fairly frequent words. However, frequency could also searve as a factor interacting with analogy. I leave this issue for future studies.

(30) Novel words used in the experiment

a. Bisyllabic forms

CV.CV.		CVC.CV.		CV.CVC.		CVC.CVC.	
NOVEL	REAL	NOVEL	REAL	NOVEL	REAL	NOVEL	REAL
suga	súdá	nɨkbe	nɨ́kdé	kɨriŋ	kɨ́rím	əkc^{h}əp	ə́kc^{h}ə́k
cop^{h}a	cók^{h}á	tamge	támbé	hobat	hóbák	pəpsət	pə́psə́k
keni	kémi	cipde	cípge	kasɨŋ	kásɨm	pancup	páncuk
nadi	nábi	kukda	kúkca	padap	pádak	tangoŋ	tángol
cagu	cadú	nemgi	nembí	kɨnɨm	kɨnɨ́l	əlgun	əlgúl
kubu	kudú	kamba	kamcá	kusət	kusə́k	hətgaŋ	hətgán

b. Trisyllabic forms

CV.CV.CV.		CVC.CV.CV.		CV.CV.CVC.		CVC.CVC.CV.	
NOVEL	REAL	NOVEL	REAL	NOVEL	REAL	NOVEL	REAL
mucide	múcíge	simmami	símmáni	ərɨsiŋ	ə́rɨ́sin	caŋttangi	cáŋttánci
sat^{h}umi	sát^{h}úri	toŋc^{h}iri	tóŋc^{h}ími	jədɨnɨm	jə́dɨ́rɨm	kombaŋge	kómbáŋde
mudəbi	mudə́gi	talgugi	talgúci	arimaŋ	aríraŋ	paŋraŋi	paŋmáŋi
toragi	toráci	tanmugi	tanmúci	ttɨgegil	ttɨgécil	incəlri	incə́lmi
sadami	sadarí	coŋami	coŋári	susuttaŋ	susukkáŋ	cindamme	cindallé
nəsɨme	nəsɨré	nalc^{h}isi	nalc^{h}ɨ́gi	kədɨrɨŋ	kədɨrɨ́m	mindɨnne	mindɨllé

5.2.3 Hypothesis and Predictions

The *Word Similarity Effect Hypothesis* predicts that the patterns of existing similar words should serve as a factor determining the NKK accent assignment of novel words. In order to test this hypothesis, the experiment employed novel words that were identical to existing words except for one consonant. If analogy to patterns of phonetically similar words crucially influences the accent placement in novel words, we expect that patterns emerging in novel words should not be significantly different from the patterns in the phonetically similar existing counterparts.

5.2.4 Results

A total of 1434 tokens (48 items x 30 subjects, excluding six tokens from the analyses)[18] were collected for the analogy test: 719 bisyllabic and 715 trisyllabic tokens. Accent patterns of novel words were compared with those of existing analogous counterparts. Two types of statistical analysis were employed: a loglinear analysis for comparison of accent patterns between the two groups (novel words vs. existing words) and Pearson's Chi-square for the main effect of syllable structure and word similarity (analogical accent type). The independent variables were (i) accent position in novel words; (ii) syllable structure, e.g. LL, HL, LH, HH (H=heavy; L=light); and

[18] Antepenultimate accent was rarely produced in novel words. Because only 5 of 720 tokens were produced with antepenultimate accent, these 5 tokens were omitted in the statistical analyses. In addition, one token was excluded due to production error.

(iii) accent position in existing counterparts. The dependent variable was the number of tokens for each accent pattern (double; penultimate; final).

5.2.4.1 Overall Results: Analogy Effects

The overall results are given in Table 11.

Table 11 Overall distribution of accent in novel words

	Agree w/ analogical patterns	Disagree w/ analogical patterns	Total
Bisyllabic	284(39%)	435(61%)	719
Trisyllabic	377(53%)	338(47%)	715
Total	661(46%)	773(54%)	1434

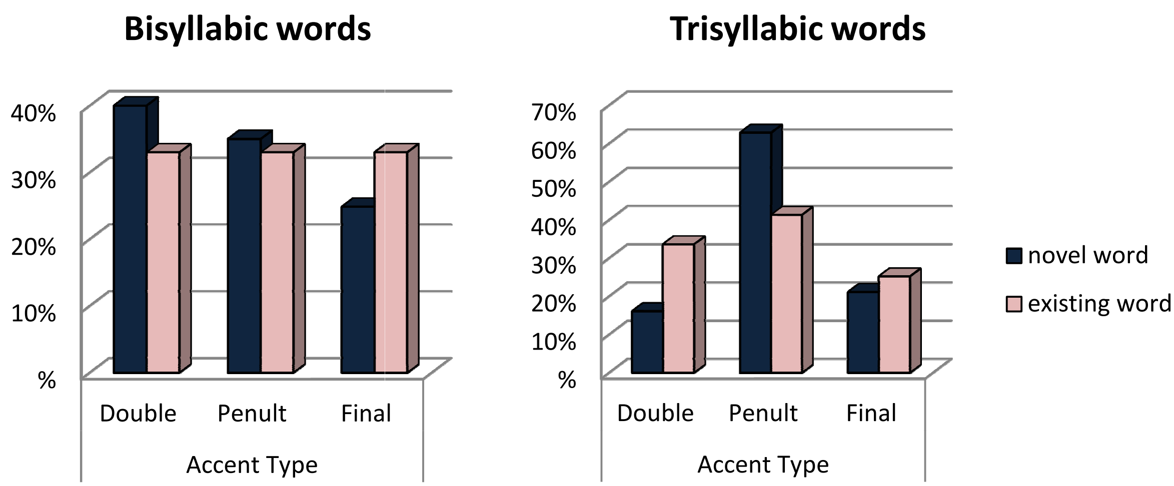
As shown in Table 11, the overall distribution of accent in novel words does not follow the distribution of accent in phonetically similar existing words. Only 46% of responses (661 out of 1434 tokens) followed analogical patterns: 39% in bisyllabic words and 53% in trisyllabic words.

In bisyllabic words, as the graphs in Figure 32 illustrate, double accent was the most common pattern (40%: 288/719) and was actually more frequent in novel words than in phonetically similar real words (33%). In contrast, final accent was less frequent in novel words (25%:179/719) than in phonetically similar real words (33%).

In trisyllabic words, penultimate accent was the most frequent pattern in novel words (63%; 449/715), whereas penultimate accent occurred in only 41% of the analogical real words. Double accent assignment also did not match the analogical patterns: 16% (115/715) of novel words were produced with double accent, as opposed to 34% of phonetically similar existing words. Final accent was the only type showing a similar frequency in novel and phonetically similar real words: final accent was found in 21% (151/715) of novel words and 25% of existing words.

A loglinear analysis confirmed that the distribution of the accent patterns in novel vs. phonetically similar existing words was significantly different ($\chi^2(3) = 84.28$, $p < .001$). It also showed that interaction between syllable structure and accent type was significant ($\chi^2(9) = 427.72$, $p < .001$).

Figure 32 Accent distribution in novel words vs. phonetically similar existing words



In the next sections, I consider the patterns in bisyllabic and trisyllabic words in greater detail.

5.2.4.2 Effects of Syllable Structure

In the previous chapters, I have shown that syllable structure played a role in accent placement in novel words which did not have obvious existing counterparts, and in statistical accent patterning in the lexicon. Both in the novel words and in the lexical statistics, the following syllable structure-sensitive tendencies were observed: (i) heavy syllables tended to attract accent in both bisyllabic and trisyllabic words: words with final heavy syllables (e.g. LH, LLH) tended to have final accent and words with initial heavy syllables (e.g. HLL, HHL) tended to have double accent; (ii) in words without heavy syllables, penultimate accent was the most frequent pattern; and (iii) in trisyllabic existing words, penultimate accent was dominant across all the word types except LLH (words with final heavy syllables). These syllable structure-sensitive tendencies held even where analogy pulled in a different direction.

The number of novel word tokens produced with each accent pattern was counted according to syllable structure combinations. The results for bisyllabic words are presented first.

Bisyllabic words

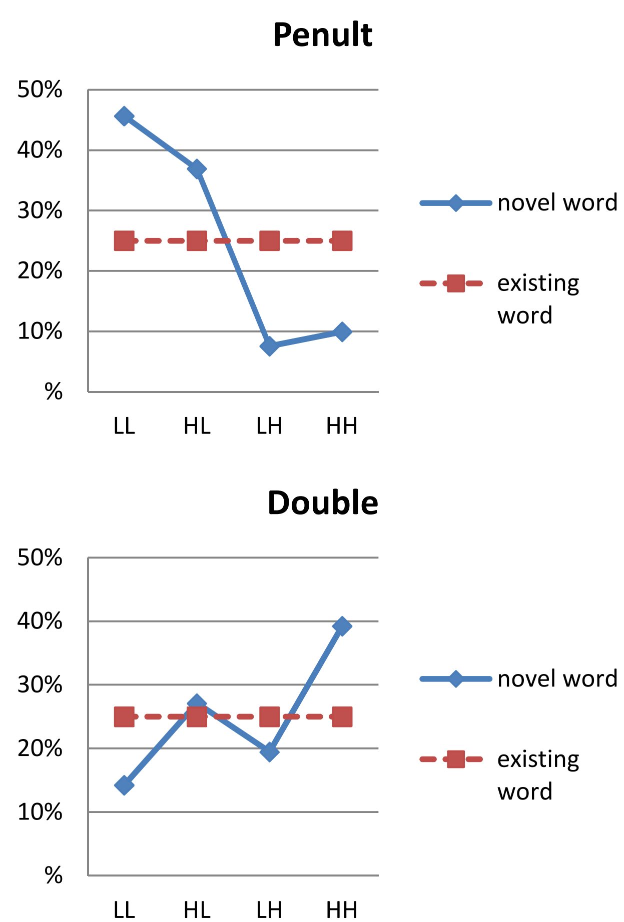
The frequency of each accent response for each word type in bisyllabic words is presented in Table 12. The number in parentheses indicates proportion of each accent, and the most frequent accent type for each word type is indicated in bold.

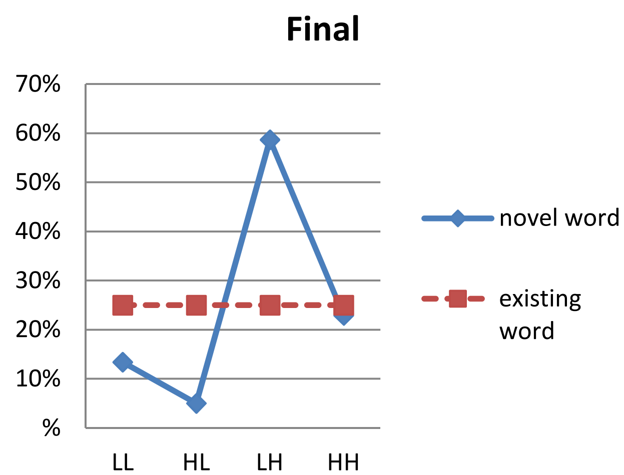
Table 12 Accent distribution according to syllable structure in bisyllabic novel words

Word Type	Accent Type			Total
	Double (%)	Penult (%)	Final (%)	
LL	41(23)	**115(64)**	24(13)	180
HL	78(43)	**93(52)**	9(5)	180
LH	56(31)	19(11)	**105(58)**	180
HH	**113(63)**	25(14)	41(23)	179
Total	288(40)	252(35)	179(25)	719

Penultimate accent was the most frequent pattern in LL and HL, while final accent and double accent were the most common patterns in LH and in HH, respectively. These results are consistent with those found in Experiment 1, and contradict the predictions of the hypothesis that speakers should accent novel words in accord with the accent pattern of phonetically similar existing words. Because the novel words were constructed so that their existing counterparts displayed equal distribution of accent type across different word types, as shown in Figure 33, analogy alone would predict equal proportions of each accent type. Contra the prediction, the accent placement of novel words displayed uneven distribution according to syllable structure, as illustrated in the following graphs in Figure 33.

Figure 33 Proportion of each accent pattern in novel words vs. existing counterparts





Penultimate accent was far more frequent in LL novel forms than in the analogous existing words, but less frequent in LH and HH novel forms. The distribution of non-penultimate accent in novel forms also did not parallel that in phonetically similar existing words. Double accent was more frequent in novel forms with initial heavy syllables (HL and HH) and final accent was more frequent in LH novel forms than in analogous real words. Therefore, as in Experiment 1, distribution was skewed toward accenting heavy syllables in novel words, even where a phonetically similar existing word had accent on a non-heavy syllable.

The main effects of syllable structure were significant for all the accent patterns in bisyllabic novel words: for penultimate accent, χ^2 (3)=169.9, p<.001, N=719; for double accent, χ^2 (3)=68.9, p<.001, N=719; and for final accent, χ^2 (3)=158.9, p<.001, N=719.

Trisyllabic words

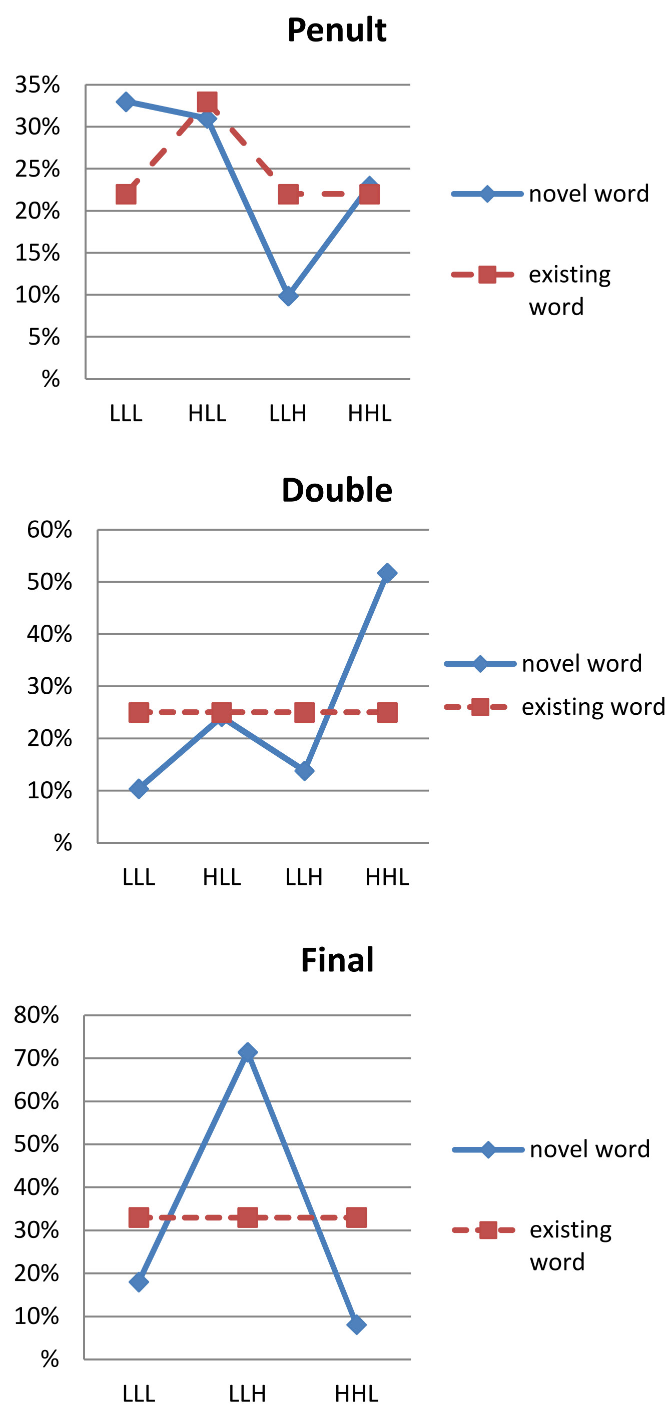
Results for trisyllabic words showed that preference for penultimate accent was fairly strong in all word types except LLH. The results are presented in Table 13.

Table 13 Distribution of accent patterns according to word type

Word Type	Accent Type			Total
	Double(%)	Penult(%)	Final(%)	
LLL	11(6)	**150(83)**	19(11)	180
HLL	28(16)	**143(82)**	4(2)	175
LLH	16(9)	49(27)	**115(64)**	180
HHL	60(33)	**107(59)**	13(7)	180
Total	115(16)	**449(63)**	151(21)	715

Despite the overall tendency toward penultimate accent, final accent was the most frequent pattern in LLH, which was consistent with the results in Experiment 1. The preference for double accent in HLL and in HHL was somewhat weaker compared to that in novel words without obvious existing counterparts in Experiment 1. Neither the predominant penultimate accent nor the strong final accent in LLH paralleled the analogical patterns of phonetically similar existing words. The following graphs in Figure 34 illustrate the mismatch between trisyllabic novel words and analogous existing words.

Figure 34 Proportion of each accent in novel words vs. existing counterparts



As shown in Figure 34, penultimate accent was far more frequent in LLL novel forms than in analogous existing words, but less frequent in LLH novel forms. Double accent was more frequent in HHL novel forms and final accent was more frequent in LLH novel forms than in analogous real words. These distributions were in accord with the syllable structure-sensitive tendencies that emerged in the previous chapters. Therefore, the analogical patterns of phonetically similar real words do not appear to serve as a major factor in the patterns of novel words.

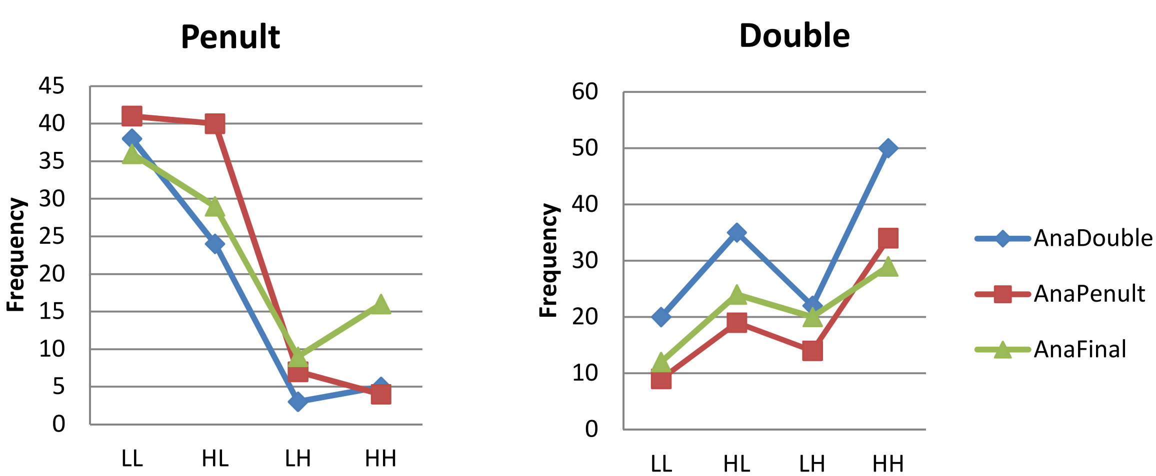
The main effects of syllable structure were significant for all the accent patterns in trisyllabic novel words: for penultimate accent, $\chi^2(3)=157.7$, $p<.001$, $N=715$; for double accent, $\chi^2(3)=59.9$, $p<.001$, $N=715$; and for final accent, $\chi^2(3)=267.8$, $p<.001$, $N=715$.

5.2.4.3 Biased analogy effects toward syllable structure-sensitive patterning

The results showed that the accentuation of novel words was affected by syllable structure, rather than the accent patterns of phonetically similar existing words. Nevertheless, effects of analogy with existing words were found in some cases. Interestingly, the analogy effects held only in the direction of the syllable structure-sensitive patterning. That is, the assignment of analogical patterns was more likely when the analogical patterns matched the syllable structure-sensitive tendencies than when the patterns were in conflict with these tendencies. In consequence, the syllable structure-sensitive tendencies in novel words were never overridden by the analogy effects, but the two factors, analogy and syllable structure sensitivity, had a cumulative effect.

The following graphs in Figure 35 illustrate the biased analogy effects in bisyllabic novel words. The chart on the left-hand side represents the frequencies of penultimate accent in novel words of different syllable type where the phonetically analogous words had one of three accent patterns. The chart on the right-hand side represents the frequencies of double accent depending on analogical accent type.

Figure 35 Raw frequency of each accent in bisyllabic novel words

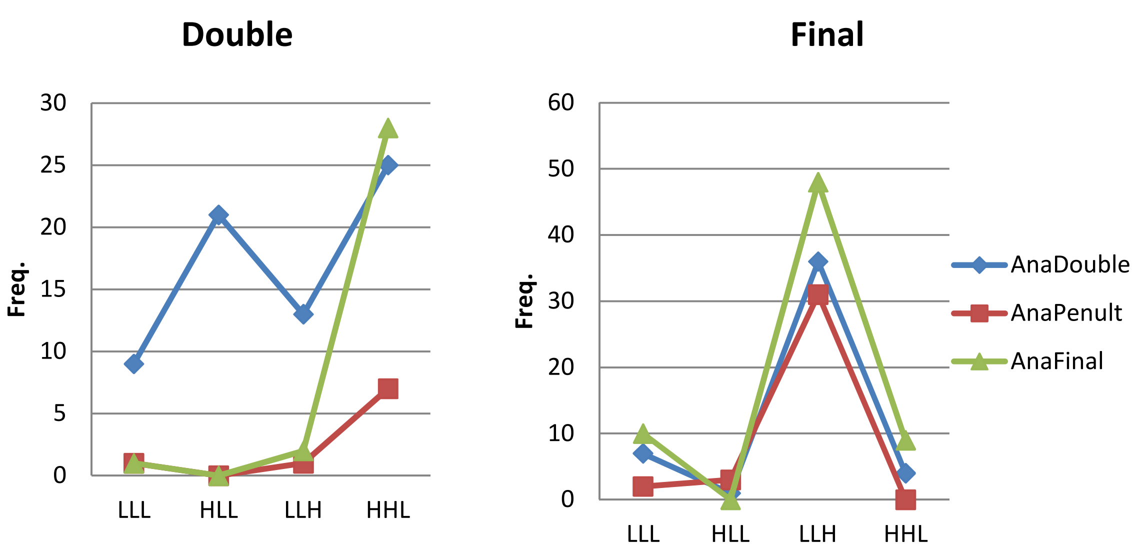


(i) AnaDouble**:** the cases when existing counterparts carried double accent; (ii) AnaPenult: the cases when existing counterparts carried penultimate accent; and (iii) AnaFinal: the cases when existing counterparts had final accent.

As shown in the chart on the left-hand side in Figure 35, in LL and HL forms, penultimate accent was more frequent when the existing counterparts carried penultimate accent (AnaPenult) than when the existing counterparts carried non-penultimate accent patterns (AnaDouble; AnaFinal). The effect of analogical accent patterns was significant in HL words ($\chi^2(2)=8.94$, p=.01). It is interesting that the analogy effects were not consistent in LH and in HH. In LH forms, the assignment of penultimate accent was not much different for forms with and without similar counterparts. For HH forms, penultimate accent was actually produced more frequently when analogy predicted *final* accent than when analogy predicted penultimate accent. On the other hand, the chart on the right-hand side in Figure 35 shows that double accent was strikingly more frequent in HL and HH (but not in LL and LH forms) when analogy predicted double accent. The effects of analogical accent patterns were significant in HL and HH ($\chi^2(2)=9.1$, p=.01 in HL; $\chi^2(2)=18.6$, p<.001 in HH).

In trisyllabic words, there were cases where accent patterns were affected by analogical accent patterns, as shown in Figure 36. However, it is also notable that the effects were not consistent across the word types.

Figure 36 Raw frequency of each accent in trisyllabic novel words

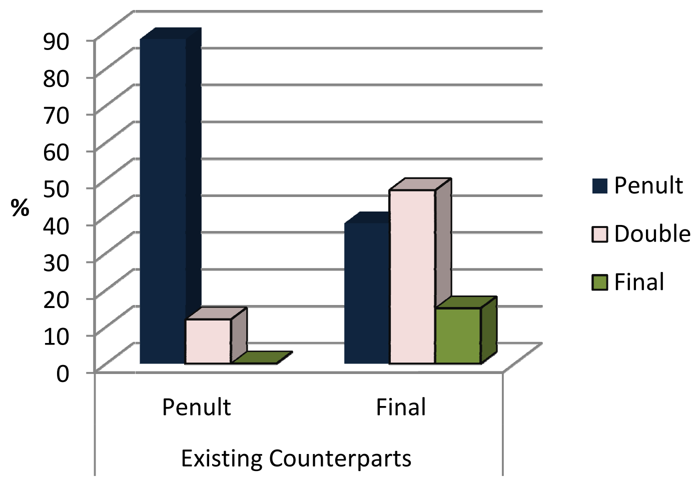


(i) AnaDouble**:** the cases when existing counterparts carried double accent; (ii) AnaPenult: the cases when existing counterparts carried penultimate accent; and (iii) AnaFinal: the cases when existing counterparts had final accent.

As shown in the graph on the left-hand side in Figure 36, double accent was more frequent when analogous existing words had double accent than when analogous existing words had penultimate accent. Notably, in HLL and HHL, double accent was strikingly higher than in other word types. The statistical analyses confirmed that the effects of analogical accent patterns were significant ($\chi^2(1)=36.6$, p<.001 in HLL; $\chi^2(2)=19.4$, p<.001 in HHL). It is intriguing, however, that double accent in HHL was far more frequent even in cases when the accent was in conflict with the analogical patterns: for example, double accent was more frequent even when the analogous existing word had *final* accent. On the other hand, the graph on the right-hand side in Figure 36 shows that final accent in LLH was even more frequent when the analogous existing word had final accent vs. other accent patterns ($\chi^2(2)=11.02$, p=.004), although the preference for final accent in LLH was strong in general regardless of analogical accent patterns. The analogy effect on final accent was notable in LLH but not in other word types.

The specific patterns of HHL forms are repeated in Figure 37, because they demonstrate the biased analogy effects toward the structure-sensitive patterning rather straightforwardly.

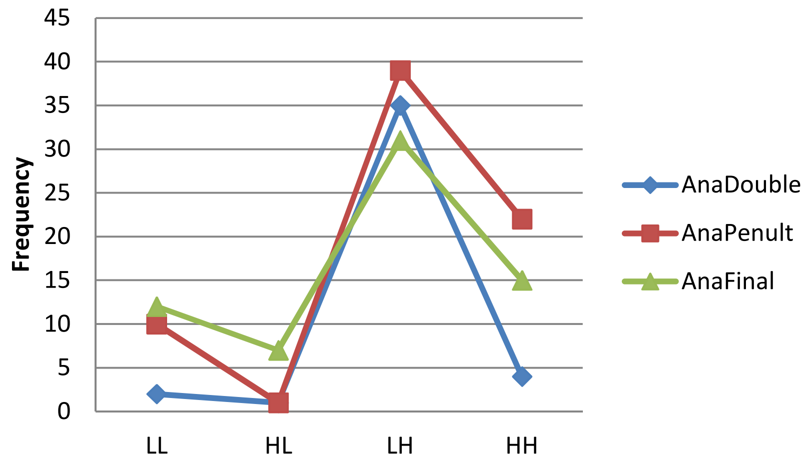
Figure 37 Accent distribution in HHL novel words according to analogical accent patterns



Recall that penultimate accent was the most frequent pattern in general and in HHL (refer to Table 13). Figure 37 shows that the preference for penultimate accent was even stronger when the existing counterparts had penultimate accent (88%), whereas only 38% of novel forms were produced with penultimate accent when the existing counterparts carried final accent. Nevertheless, final accent in HHL existing counterparts did not reinforce the assignment of final accent in HHL novel forms, either: only 15% of HHL novel forms were produced with final accent. This shows that the analogy effects did not hold when the analogical patterns contradicted the structure-sensitive tendencies. Furthermore, double accent appeared more frequently than penultimate accent (47% vs. 38%) in HHL when the existing counterparts had final accent. This indicates that penultimate accent yielded to double accent in HHL when the pattern was not associated with analogy.

I have shown that analogical patterns of existing words were relevant when the patterns were in accordance with the structure-sensitive tendencies. However, the analogy effects did not operate for the patterns which contradicted the structure-sensitive tendencies. Furthermore, there was a case in which analogy with existing words was not involved at all even for the patterns which were consistent with the structure-sensitive tendencies: assignment of final accent in bisyllabic words was totally independent of accent patterns in analogous existing words.

Figure 38 No analogy interference in the assignment of final accent in bisyllabic novel words



(i) AnaDouble**:** the cases when existing counterparts carried double accent; (ii) AnaPenult: the cases when existing counterparts carried penultimate accent; and (iii) AnaFinal: the cases when existing counterparts had final accent.

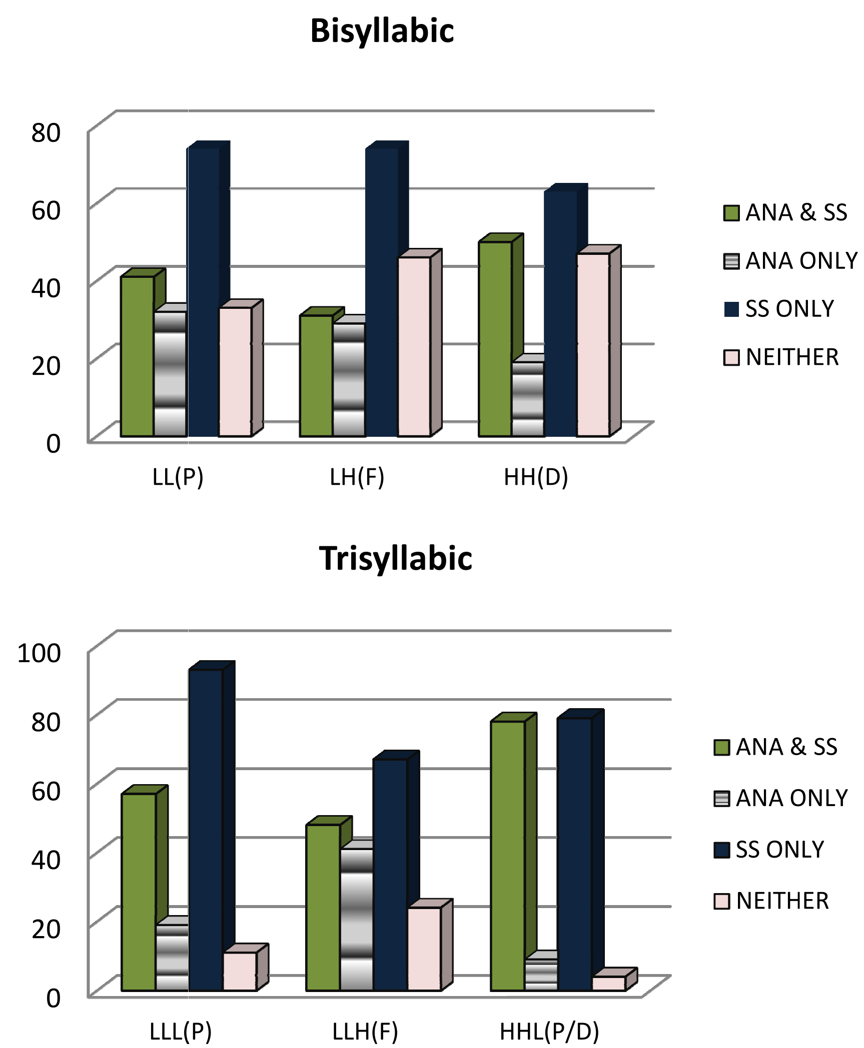
As shown in Figure 38, the frequency of final accent was not necessarily boosted by the homogeneous accent pattern in the analogous existing counterparts. For example, final accent in LH was produced *less* frequently when the analogous existing words had final accent than when the analogous existing words had non-final accent (31% in AnaFinal vs. 35% in AnaDouble vs. 39% in AnaPenult). Likewise, final accent placement in HH also did not show effects of the analogical accent patterns: final accent was more frequent when the analogous existing words had penultimate accent. These patterns demonstrate that final accent in bisyllabic novel words was assigned independently of the patterns of phonetically similar existing words.

To summarize, accent patterns in novel words were significantly different from the patterns in analogous existing counterparts. On the other hand, syllable structure played a crucial role in accent assignment of novel words with analogous counterparts, as in Experiment 1 with novel words without obvious existing counterparts. The effect of syllable structure was independent of the analogical patterns because the existing counterparts displayed equal distribution of accent type across different word types. However, the frequency of the analogical patterns increased when analogical patterns matched the syllable structure-sensitive tendencies. Hence, the analogy effects were also dependent on syllable structure. Taken all together, analogy to patterns of phonetically similar existing words alone cannot explain NKK accent patterns in novel words. The biased analogy effects and their implications are discussed in the following section.

5.3 Discussion

The previous section showed that the *Word Similarity Effect Hypothesis* did not hold in NKK accent patterns in novel words. Rather, NKK speakers accented the novel words in accord with the structure-sensitive tendencies. The syllable structure sensitivity was stronger than analogy even in the novel forms where the analogical patterns contradicted the structure-sensitive tendencies. The stronger effect of syllable structure is illustrated graphically in Figure 39:

Figure 39 Patterns following analogy vs. Patterns following syllable structure



(i) ANA&SS: Responses following both analogy and syllable structure; (ii) ANA ONLY: Responses following analogy rather than syllable structure; (iii) SS ONLY: Responses following syllable structure rather than analogy; (iv) NEITHER: Responses that follow neither.
*As for responses following syllable structure, I counted responses with penultimate accent in LL/LLL; responses with final accent in LH/LLH; responses with double accent in HH; and responses with double/ penultimate accent in HHL.

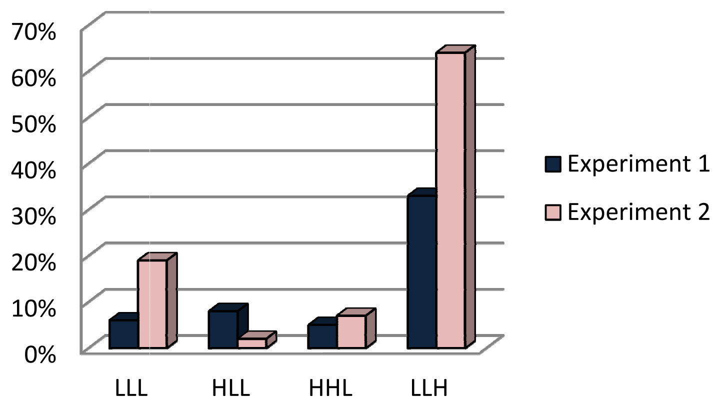
In Figure 39, the results were categorized into four different groups: (i) responses following both analogy and syllable structure; (ii) responses following analogy rather than syllable structure; (iii) responses following syllable structure rather than analogy; (iv) responses that follow neither. As shown in Figure 39, patterns following syllable structure rather than analogy (SS ONLY responses) were most frequent across all the word types both in bisyllabic and in trisyllabic novel forms, followed by patterns following both analogy and syllable structure (ANA&SS responses). In contrast, responses following analogy rather than syllable structure (ANA ONLY responses) were least frequent in bisyllabic novel forms and less frequent than SS ONLY and ANA&SS responses in trisyllabic novel forms. This clearly indicates that the effects of syllable structure outweigh analogy in NKK accent patterns in novel words.

The analogy effects were notable only when analogical accent patterns were in accordance with the syllable structure-sensitive tendencies, in that the assignment of analogical patterns was more likely when the analogical patterns matched the syllable structure-sensitive tendencies than when the patterns were in conflict with these tendencies. For example, the preference for final accent in LLH was greater when the analogical words had final accent. In contrast, the patterns which contradicted the structure-sensitive tendencies (e.g. penultimate accent; double accent in LLH) were not necessarily boosted by the analogical patterns.

The tendency toward accenting heavy syllables parallels the patterns found in novel words which do not have obvious existing counterparts,[19] as shown in Experiment 1. However, the syllable structure-sensitive tendencies appeared even more strongly in novel forms which were analogous to real words than in novel words lacking obvious existing counterparts. The following graph in Figure 40 illustrates the different assignment of final accent in novel forms with no obvious phonetically similar counterparts (Experiment 1) vs. novel forms with existing similar forms with final accent (Experiment 2).

[19] The novel forms used in Experiment 1 were constructed using a single vowel, either *a* or *i*, for all syllables within a word; for example, *takapa* or *citiki*. Furthermore, each word contained only obstruent onsets such as *p, t, k* (e.g. *kapa*) or only sonorant onsets such as *m, n* (e.g. *mana*). Therefore, the combination of a single vowel and the restricted onset consonants made the novel words relatively less similar to existing words than the novel words used in Experiment 2, which were created based on existing words. Most novel words used in Experiment1 lack of *obviously* phonetically similar existing counterparts: i) there are only a few cases in bisyllabic words which might have analogous existing words, which are a product brand name or differ in syllable structure: e.g. novel form k^hap^ha ~ a brand name k^hap^ha 'kappa'; novel form *kaba* ~ existing word *kabaŋ* 'bag'; ii) for trisyllabic novel forms, existing counterparts are very unlikely overall.

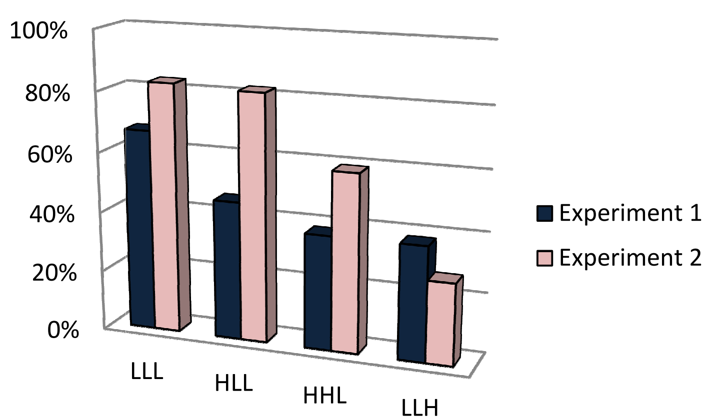
Figure 40 Proportion of final accent in trisyllabic novel words: Experiment 1 vs. Experiment 2



Final accent was far more frequent (64%) in novel words with existing similar final accented LLH words (Experiment 2) than in those lacking similar LLH counterparts (33%) (Experiment1). The higher frequency of final accent in LLH novel forms in Experiment 2 is attributable to the biased analogy effects, which reinforced accenting final heavy syllables: the final accent reinforcement by analogy was not consistent across other word types, although they had final accented existing counterparts. Therefore, the stronger preference for final accented LLH words in Experiment 2 compared to that in Experiment 1 verifies that the analogy effect was indeed limited for the syllable structure-sensitive patterning.

In addition, penultimate accent also appeared stronger in Experiment 2 than in Experiment 1. Penultimate accent was the most common pattern in trisyllabic novel forms in Experiment 1 as well as in Experiment 2, but the pattern was even stronger when reinforced by analogy. Figure 41 presents the comparison of the occurrence of penultimate accent between Experiment 1 and Experiment 2.

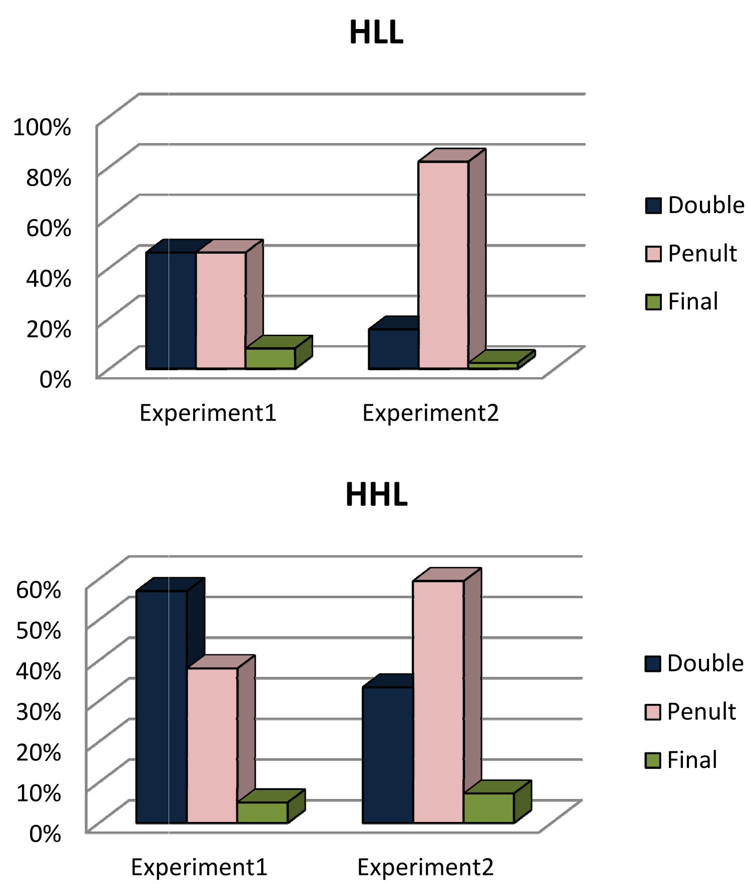
Figure 41 Proportion of penultimate accent in trisyllabic novel words: Experiment 1 vs. Experiment 2



As shown in Figure 41, penultimate accent was more frequent in Experiment 2 than in Experiment 1, and this tendency was consistent across all the word types except LLH. The stronger preference for penultimate accent in Experiment 2 was attributable to the effects of analogy with existing similar penultimate accented words. However, the strengthening of penultimate accent by analogy disappeared in LLH because penultimate accent in LLH contradicted the structure-sensitive tendencies. Thus, the analogy effects emerging in the assignment of penultimate accent were also sensitive to syllable structure.

These results raise the question of how to explain the stronger preference shown in Experiment 2 for penultimate accent in HLL and HHL even though this preference did not follow the overall tendency towards accenting heavy syllables. Recall that in Experiment 1, double accent was fairly common in HLL and even more frequent than penultimate accent in HHL, as shown in Figure 42, consistent with the structure-sensitive tendencies. However, in Experiment 2 where HHL/HLL novel forms had analogous existing counterparts, penultimate accent overrode double accent. The charts in Figure 42 illustrate the different accent distribution in HLL/HHL novel forms in Experiment 1 vs. Experiment 2.

Figure 42 Accent distribution in HLL/HHL: Experiment 1 vs. Experiment 2



These results suggest that NKK speakers are more likely to choose penultimate accent over double accent for trisyllabic novel forms with initial heavy syllables in cases where analogy motivates penultimate accent. Consequently, double accent was assigned less frequently in Experiment 2 than in Experiment 1. Furthermore, penultimate syllables are the generally preferred accent position in existing trisyllabic words and the analogy effects with existing words appeared more strongly in the direction toward the preferred position than toward accented heavy syllables.

5.4 Remark on Competing Analogous Counterparts

One possible confounding factor is that some of the bisyllabic novel forms in Experiment 2 might have competing analogous existing counterparts with different accent patterns. For example, for a novel form *cagu*, the real word *cigu* 'Earth' (with double accent) differs from the corresponding novel form in its initial vowel rather than in its final consonant, and therefore might compete with the existing counterpart *cadu* 'plum' (with final accent). Thus, in this section I considered accent patterns of competing analogous counterparts.

I collected additional analogous existing words which might compete with the base counterparts listed in (31), by asking three NKK informants to write, for each novel form of the stimuli used in Experiment 2, as many existing words as each novel form brought to mind. For most of the bisyllabic novel forms, one or two other existing words in addition to the base counterparts were produced, except for three novel forms, *nɨkbe*, *əkc^{h}əp*, *əlgun*, where only the base counterpart was available. As shown in (31), I categorized the additional analogous counterparts[20] into three different groups: i) a nearly-identical group (identical to a novel form except for some other single segment); e.g. a real word *cɨ́gú* 'Earth' vs. a novel form *cagu*; ii) a similar group (identical to a novel form except for two segments or share an identical final syllable); e.g. a real word *kkɨ́ni* 'meal' vs. a novel form *keni*; and iii) a nearly-identical group but with different syllable structure (NI&diff.SS) (identical to a novel form except for one segment but differs in syllable structure); e.g. a real word *cə́guk* 'trace' vs. a novel form *cagu*.

(31) Additional analogous existing counterparts

(i) Nearly-Identical group		(ii) Similar group		(iii) NI&diff.SS group	
Existing	*Novel*	*Existing*	*Novel*	*Existing*	*Novel*
cígú	cagu	kkíni	keni	súgap	suga
cup^{h}á	cop^{h}a	ssɨ́lgé	tamge	cáguk	cagu
tubú	kubu	pindé	cipde	cóp^{h}án	cop^{h}a
madí	nadi	kjésɨ́ŋ	kasɨŋ	kúbun	kubu
kúksá	kukda	tédáp	padap	kámpáŋ	kamba
néŋgí	nemgi	kɨ́nóm	kɨnɨm	hə́tge	hətgaŋ
kɨríl	kɨriŋ	pə́sət	kusət	pə́sət	pəpsət
hobán	hobat	seŋgáŋ	hətgaŋ		
		tangún	tangoŋ		
		pánccók	pancup		

[20] Non-phonetically related responses were excluded from the analyses.

The number of tokens consistent with analogical accent patterns was counted for each group. The results are presented in the following table:

Table 14 Frequency of tokens following analogical patterns[21]

Agree with Near-Idnt group	Agree with Similar group	Agree with NI&Diff.SS group
69/240(29%)	107/300(35%)	104/210(50%)

As shown in Table 14, the accent patterns in bisyllabic novel forms did not consistently follow analogical patterns for either of the competing analogous groups, a nearly-identical or a similar group. Recall that 39% of the responses (284/720) followed accent patterns of the base-existing counterparts in Experiment 2. The accent patterns of the competing counterparts were actually followed less frequently than those of the counterparts assumed in the original analysis. As for NI&Diff.SS group, the proportion of tokens following analogical patterns was relatively higher in this group than in other groups (50% vs. 29%, 33%). However, in some cases, it is not clear whether analogy or syllable structure triggered the analogical patterns because syllable structure in the existing counterparts of this group is different from the corresponding novel forms, and the patterns were consistent with the syllable structure-sensitive patterns (e.g. analogical penultimate accent for a novel form *súga*, c.f. a real word *súgap* 'handcuff'). Therefore, analogy is not an exclusive factor of the patterns in NI&Diff.SS group, so the results of this group will not be considered any further. Nevertheless, all these together suggest that NKK speakers were not likely to follow analogical patterns even in the cases when they came up with alternative analogical words. Therefore, the overall results remain intact although other competing analogous counterparts were present.

Unlike bisyllabic novel forms, trisyllabic novel forms rarely had additional analogous counterparts. In the Korean lexicon, neighboring lexical words which share an identical onset in two or more syllables are not common. Actually, all three informants gave only the predicted analogous base words for 21 of the 24 trisyllabic novel forms.[22] Therefore, the trisyllabic novel forms in Experiment 2 almost overwhelmingly reflect the target existing counterparts.

5.5 Conclusion

This chapter investigated analogy effects on NKK accent assignment of novel words, testing the *Word Similarity Effect Hypothesis*: analogy to patterns of phonetically similar words is a major

[21] The results of an individual item are presented in Table 17 in Appendix III.

[22] Only three items of the 24 trisyllabic novel forms were given additional existing counterparts: *mudə́wi* 'sweltering heat'(c.f. base-counterpart *mudə́gi* 'pile') for a novel form *mudəbi*; *sasími* 'a loanword from Japanese, meaning raw fish' (c.f. base-counterpart *sadarí* 'ladder') for a novel form *sadami*; and *báŋráŋca* 'wanderer' (c.f. base-counterpart *baŋmáŋi* 'cudgel') for a novel form *baŋraŋi*. The competing accent patterns of these alternative analogous words were consistent with the syllable structure-sensitive tendencies (penultimate accent in words without heavy syllables; double accent in words beginning with a heavy syllable). Therefore, the presence of the competing analogous counterparts in trisyllabic novel forms should not be considered meaningful where analogical patterns compete with structure-sensitive patterning.

factor in NKK accentuation of novel words. The results showed that word similarity did not play an important role in accent placement of novel words. Rather, NKK speakers accented novel words in accordance with the syllable structure-sensitive tendencies, even when accent patterns in analogous existing words contradicted the syllable structure-sensitive tendencies. This indicates that structure sensitivity was stronger than analogy in NKK accentuation of novel words. There were cases where accent patterns of analogous existing words were relevant to accent patterns in novel forms, but even the analogy effects were not independent of syllable structure: the analogy effects were biased toward the structure-sensitive patterning. The biased analogy effects resulted in stronger tendencies toward penultimate accented trisyllabic forms and accented final heavy syllables in novel words with obvious existing counterparts than in novel forms without them. I conclude that NKK accent patterns of novel words indeed represent NKK speakers' abstract phonological knowledge rather than a mere product of analogy, and that their knowledge of relationship between syllable structure and accent also constrains the effects of analogy.

Appendix III: Results of Experiment 2

Table 15 Results in bisyllabic words in Experiment 2

Novel words	Similar existing words	Analogical pattern	Accent in novel words			Total
			Double	Penult	Final	
suga	suda	double	7	23	0	30
cagu	cadu	final	6	14	10	30
cop^ha	cok^ha	double	13	15	2	30
kubu	kudu	final	6	22	2	30
keni	kemi	penult	6	23	1	30
nadi	nabi	penult	3	18	9	30
nɨkbe	nɨkde	double	18	11	1	30
tamge	tambe	double	17	13	0	30
cipde	cipge	penult	13	16	1	30
kukda	kukca	penult	6	24	0	30
nemgi	nembi	final	9	14	7	30
kamba	kamca	final	15	15	0	30
kasɨŋ	kasɨm	penult	10	4	16	30
padap	padak	penult	4	3	23	30
kɨriŋ	kɨrim	double	7	0	23	30
hobat	hobak	double	15	3	12	30
kɨnɨm	kɨnɨl	final	10	1	19	30
kusət	kusək	final	10	8	12	30
$əkc^həp$	$əkc^hək$	double	25	3	2	30
pəpsət	pəpsək	double	25	2	2	29
hətgaŋ	hətgan	final	14	15	1	30
əlgun	əlgul	final	15	1	14	30
tangoŋ	tangol	penult	15	1	14	30
pancup	pancuk	penult	19	3	8	30
Total			288	252	179	719

Table 16 Results in trisyllabic words in Experiment 2

Novel words	Similar existing words	Analogy Accent	Accent in novel words			Total
			Double	**Penult**	**Final**	
sat^{h}umi	sat^{h}uri	double	5	23	2	30
mucide	mucige	double	4	21	5	30
mudəbi	mudəgi	penult	1	27	2	30
toragi	toraci	penult	0	30	0	30
sadami	sadari	final	1	27	2	30
nəsɨme	nəsɨre	final	0	22	8	30
nalc^{h}isi	nalc^{h}igi	penult	7	18	0	25
simmami	simmani	double	10	20	0	30
toŋc^{h}iri	toŋc^{h}imi	double	11	18	1	30
talgugi	talguci	penult	0	29	1	30
tanmugi	tanmuci	penult	0	29	1	30
coŋami	coŋari	penult	0	29	1	30
jədɨnɨm	jədɨrɨm	double	2	8	20	30
ərɨsiŋ	ərɨsin	double	11	3	16	30
arimaŋ	ariraŋ	penult	0	7	23	30
susuttaŋ	susukkaŋ	final	0	5	25	30
ttɨgegil	ttɨgecil	penult	1	21	8	30
kədɨrɨŋ	kədɨrɨm	final	2	5	23	30
kombaŋge	kombaŋde	double	11	16	3	30
caŋttaŋgi	caŋttanci	double	14	15	1	30
incəlri	incəlmi	penult	2	28	0	30
mindɨnne	mindɨlle	final	10	14	6	30
cindamme	cindalle	final	18	9	3	30
paŋraŋi	paŋmaŋi	penult	5	25	0	30
Total			115	449	151	715

Table 17 Neighboring analogous existing counterparts and frequency of analogical patterns

Novel words	Analogous existing words				Frequency of accent patterns			
	Base-existing words	Near-Idnt	Similar	NI&Diff. SS	Agree w/ Base	Agree w/ Near-Idnt	Agree w/ Similar	Agree w/ NI&Diff.SS
suga	súdá			súgap	7			23
cagu	cadú	cígú		cáguk	10	6		14
cop^ha	$cók^há$	$cup^há$		$cóp^hán$	13	2		13
kubu	kudú	tubú		kúbun	2	2		22
keni	kémi		kkɨ́ni		23		23	
nadi	nábi	madɨ́			18	9		
nɨkbe	nɨ́kdé				18			
tamge	támbé		ssɨ́lgé		17		17	
cipde	cípge		pindé		16		1	
kukda	kúkca	kúksá			24	6		
nemgi	nembɨ́	néŋgí			7	9		
kamba	kamcá			kámpáŋ	0			15
kasɨŋ	kásɨm		kjésɨ́ŋ		4		10	
padap	pádak		tédáp		3		4	
kɨriŋ	kɨ́rím	kɨrɨ́l			7	23		
hobat	hóbák	hobán			15	12		
kɨnɨm	kɨnɨ́l		kɨ́nóm		19		10	
kusət	kusə́k		pə́sət		12		8	
$əkc^həp$	$ə́kc^hə́k$				25			
pəpsət	pə́psə́k			pə́sət	25			2
hətgaŋ	hətgán		seŋgáŋ	hə́tge	1		1	15
əlgun	əlgúl				14			
tangoŋ	tángol		tangún		1		14	
pancup	páncuk		pánccók		3		19	
Total number of tokens with analogical patterns					284/720	69/240	107/300	104/210
%					39%	29%	35%	50%

Chapter 6 The Grammar of the NKK Accent Patterns in Novel Words

6.1 Introduction

The previous chapters have presented evidence that NKK speakers accent novel words not randomly but in a patterned manner. The results from the experimental studies in Chapter 3 and the study of lexical statistical patterns in Chapter 4 demonstrate that syllable structure plays an important role in NKK accent patterns and that the syllable structure-sensitive tendencies emerging in novel words are actually consistent with the native lexical statistical patterning. These results suggest that NKK speakers indeed have knowledge of the association of syllable structure and accent reflected in the native lexicon and apply this knowledge to novel words. Chapter 5 shows that the accent patterns of novel words are not necessarily influenced by the accent pattern of the phonetically closest existing words. Instead, NKK speakers tend to accent novel words in accordance with syllable structure, even when an analogous existing word with the same syllable structure has a different accent pattern. All these results together indicate that syllable structure is indeed important in accentuation of novel words, which suggests that the accent patterns in novel words are guided by a grammar.

Several previous studies of lexical accent patterns in NKK have assumed that there is a group of words that lack an underlying accent, called an accentless word group (e.g. Y-H Chung 1991; N-J Kim 1997; S-H Kim 1999).[23] For this accentless group, N-J Kim (1997) proposes a default penultimate accent in words containing no long vowels, based on the fact that penultimate accent is dominant in native words longer than three syllables and in loanword accentual patterns. However, S-H Kim (1999) assumes that unaccented words receive accent on the final syllable (cf. Idsardi & Chang (2003) for another approach to resolve these two conflicting arguments). However, both approaches fail to predict the syllable structure effects on accent patterns emerging in novel words, which are assumed to have no underlying accent.

I propose a stochastic grammar, constructed on the basis of native lexical patterns (e.g. Zuraw 2000; Hayes and Londe 2006), to provide an explanation of the syllable structure-sensitive tendencies emerging in the novel words. In addition, I argue that speakers learn this grammar via exposure to the lexicon. This chapter will provide a formal account of the accent patterns in novel words assuming the Gradual Learning Algorithm (GLA)(Boersma1997; Boersma & Hayes 2001). I will begin by developing a set of constraints and rankings to predict the accent patterns found in the data. Since outcomes on novel forms vary, no single ranking will derive all the forms in the data, so after developing a grammar to derive the most common patterns, I will use stochastic ranking to yield variation in accent patterns.

---

[23] Among previous studies, N-J Kim (1997) separated lexical words into three different groups according to accent patterns: words with antepenultimate accent and final accent belong to a “prelinked tone” group of which a high tone is prelinked in the underlying representation; words with double accent belong to a “floating tone” group where a high tone is floating in the underlying representation; words with penultimate accent belong to a “default tone” group where no high tone is present. He argues that a high tone would fall on the penultimate syllable when there is no high tone in the underlying representation (UR).

6.2 Optimality Theory

6.2.1 The initial state

In Optimality Theory (Prince and Smolensky 1993), the phonological grammar is modeled by a set of ranked constraints. It has been assumed that the constraints are universal, whereas the ranking differs from language to language. Therefore, the universal constraints are not necessarily learned from experience. In contrast, the language-particular ranking must be learned.

It is assumed that the universal OT constraints are innate and thus present in the initial state of the grammar. Furthermore, learnability considerations provide an argument for the claim that markedness constraints dominate faithfulness constraints at the initial state of the grammar (Smolensky 1996, Hayes 2004, Prince and Tesar 2004, Davidson *et al*. 2004). The markedness-over-faithfulness initial ranking has been supported by acquisition data in child phonology (Demuth 1995, Gnanadesikan 2004, Levelt 1995, Levelt and Van de Vijver 2004, Pater and Paradis 1996). With the markedness-over-faithfulness initial ranking, only unmarked forms will surface at the initial stage in the child's production. The process of acquisition involves promoting faithfulness constraints to approach more closely the adult grammar.

The markedness-over-faithfulness initial ranking assumes that the bias toward high-ranked markedness constraints will persist throughout learning or even in the final state of the grammar if no evidence is available to the contrary. Evidence for the persistence of markedness-over-faithfulness rankings has been proposed based on data from loanword adaptation (e.g. Shinohara 1997, 2004; Davidson *et al*. 2004) and second language acquisition (e.g. Broselow *et al*. 1998) as well as child production (e.g. Demuth 1995, Gnanadesikan 2004, Pater and Paradis 1996). For example, Shinohara (1997, 2004) shows that in the adaptation of French words into Japanese, accent patterns are fairly regular in that the antepenult is accented if the penult is light; otherwise, the penult is accented, despite unpredictable lexical accent patterns in Japanese native words. Shinohara suggests that this pattern demonstrates the effects of markedness constraints, which are uniquely apparent in loanword phonology. In addition, Shinohara suggests that some rankings among the markedness constraints emerging in loanword adaptation may be universal and derivable from the pre-fixed initial ranking of the grammar. However, Kubozono (2006) claims that the regular accent patterns emerging in loanwords actually originate from the statistics of the native lexicon and that the default antepenultimate accent reflects the most common pattern in accented native words.[24]

6.2.2 The final state

NKK accent patterns in novel words shed light on the final ranking among hidden markedness constraints which are uniquely active in novel words, since a high-ranked faithfulness constraint is not relevant in accentuation of novel words. If the language does not provide evidence for re-ranking of the markedness constraints at the initial state, the final state of the grammar will maintain the ranking at the initial stage of acquisition. However, this dissertation demonstrates that structure-sensitive accent tendencies emerging in novel words are actually consistent with the statistical patterning in the lexicon, which indicates that evidence for rankings among the

[24] Thanks to Michael Kenstowicz for drawing my attention to Kubozono's findings on accent patterns in Japanese loanwords.

markedness constraints is actually available via the lexical frequency patterns in this language. Therefore, the initial ranking among the markedness constraints will not remain intact throughout learning but will be adjusted in accordance with the lexical data. Furthermore, the NKK accent patterns emerging in novel forms are not categorical. I argue that this is because the markedness constraints are not ranked strictly at the final state. Rather, the ranking among these markedness constraints is stochastic, reflecting the frequency of existing patterns in the lexicon, which suggests that NKK learners re-ranks constraints gradually as they encounter counterevidence in the lexicon (e.g. Boersma & Hayes 2001). Although the final state of the grammar does not maintain the initial ranking identically, the effects of the initial ranking are sustained in that NKK speakers are biased toward the initial ranking among the markedness constraints, resulting in stronger regular tendencies in the accent patterns of novel words than in the patterns of existing words.

6.3 Constraints

6.3.1 Predominant penultimate accent

In the preceding chapters, we found that penultimate accent was the most common pattern overall in trisyllabic existing and novel words, particularly in words consisting only of CV syllables.[25] Recall that more than 80% of trisyllabic words consisting of only light syllables were assigned penultimate accent: 81% (88/108) of novel words and 82% (138/168) of existing trisyllabic words. This is consistent with the default penultimate accentuation, proposed by N-J Kim (1997). N-J Kim suggested the following constraints for the default penultimate accent in words containing no long vowels.[26]

(32) Constraints for default penultimate accent (based on N-J Kim 1997)[27]
..H..]$_{PW}$: A prosodic word must have a high tone.
ALIGN-R: Align the right edge of a high tone with the right edge of a word.
NONFINALITY : No high tone on the final syllable of a word is allowed.
*H: No syllable with a high tone is allowed on the surface form.

NKK is assumed to be a pitch accent language and accented syllables are always realized with a

[25] This chapter covers only trisyllabic words for theoretic analyses because bisyllabic words do not clearly demonstrate default penultimate accentuation: penultimate accent in bisyllabic words can be initial accent as well.

[26] N-J Kim argues that a default high tone would fall on the penultimate syllable when a word does not contain a long vowel. If word-final syllables contain a long vowel, however, a default high tone would fall on the final syllable. In addition to the constraints in (32), he includes the constraint WEIGHT-TO-TONE PRINCIPLE (WTP) to the default accent constraints, which forces syllables with a long vowel to be accented since coda consonants do not contribute to syllable weight under his analysis. I will add WTP to the constraint set later in this chapter when dealing with novel words which contain CVC syllables. For now, this constraint is irrelevant to the patterns in novel words with CV syllables only since novel words used in the experiments contained no long vowels.

[27] N-J Kim uses *TONELESS instead of ..H..]$_{PW}$ and DEP(ASSOC) instead of*H. ..H..]$_{PW}$ is more explicit than *TONELESS because *TONELESS impresses ambiguity between high tone and low tone although it prohibits a high toneless prosodic word in the definition. DEP(ASSOCIATION), (or NO SPREAD in Y-H Chung (2006)) has been suggested as a constraint to operate against tone spreading: An association in the output must have a correspondent in the input (N-J Kim 1997, 67). I use the surface-oriented markedness constraint *H than the ones based on covert underlying structure, which operates for the same effect conspiring with other constraints, which will be discussed later in this section.

high tone on the surface form. With the undominated constraint H]$_{PW}$, NKK prosodic words must have at least one high tone on the surface form. The default accent position is explained by the ranking of the constraints presented in (33).

(33) H]$_{PW}$ >> NONFINALITY >> ALIGN-R>>*H

The assignment of penultimate accent is illustrated in the following tableau.

(34) Penultimate accentuation

Input: |takapa|	H]$_{PW}$	NONFINALITY	ALIGN-R	*H
a. takapa	*!			
b. takapá		*!		*
c. tákapa			**!	*
d. tákápa			*	*!*
☞ e. takápa			*	*

Novel words are assumed to have no underlying high tone accent. Accent must be assigned to satisfy the constraint H]$_{PW}$. The constraint NONFINALITY, which outranks ALIGN-R, rules out the final accented candidate (34b). Antepenultimate accent (34c) loses because it violates ALIGN-R twice. Double accent (34d) is ruled out due to multiple violations of *H, leaving the penultimate accented candidate (34e) the winner.

6.3.2 Coda consonant effects

The analysis sketched in the preceding section suggests that penultimate accent should be assigned to novel forms, regardless of their syllable structure. However, although penultimate accent was the most common pattern in trisyllabic words, penultimate accent was less likely when a word contained CVC syllables word-finally or word-initially. Instead, non-penultimate accent patterns, final accent or double accent, tended to be more frequent both in existing words and in novel words, as shown in Chapters 3 and 4: words containing word-initial CVC syllables tended to exhibit double accent (e.g. *táp.ká.pa.*; *táp.kán.pa.*), while words containing final CVC syllables tended to exhibit final accent (e.g. *ta.ka.pák.*). These patterns suggest that coda consonants indeed play a role in accent assignment in novel words, and the statistical patterns in the native lexicon support this. Assuming that CVC syllables count as heavy, the structure-sensitive patterns parallel the native accent patterns where word-final/initial syllables with a long vowel (CVV) always attract accent (e.g. *maŋné:* 'the youngest'; *hɔ́:ráŋi* 'tiger') (e.g. Y-H Chung 1991; N-J Kim 1997). (See Chapter 2 for more discussion on weight-sensitive accentuation in native words.) The constraint WEIGHT-TO-TONE PRINCIPLE (WTP),[28] which forces a heavy syllable to have a high tone, should be undominated in NKK, given the fact that CVV syllables always receive a high tone. Thus, non-penultimate accent may be triggered by the constraint WTP. However, a question arises: why do CVV syllables always attract accent, while CVC syllables only show a tendency to attract accent?

[28] The constraint WTP was a revised version of WEIGHT-TO-STRESS PRINCIPLE, proposed in Prince and Smolensky (1993). The definition of WTP is repeated from Chapter 2: WEIGHT-TO-TONE PRINCIPLE (WTP): Heavy syllables must have a high tone (N-J.Kim1997; Y-H.Chung 2006; D-M.Lee 2008b).

In terms of syllable weight, CVV syllables and CVC syllables are often treated differently in many languages. For example, in Malayalam, where stress assignment is sensitive to syllable weight, CVV syllables always attract stress, while CVC syllables do not necessarily attract stress, patterning with CV syllables: that is, only syllables containing a long vowel count as heavy (Broselow *et al*. 1997). In contrast, in some languages (e.g. Yana, Finnish, Hindi), CVC syllables pattern with CVV syllables in that CVC syllables are considered heavy like CVV syllables (Broselow *et al*. 1997; Gordon 1999, 2004). For example, Hindi has been assumed to have a three-way contrast in syllable weight: light (CV), heavy (CVV or CVC) and superheavy (CVVC or CVCC), and Hindi stress patterns demonstrate that coda consonants always contribute to syllable weight: stress falls on a superheavy syllable (CVVC or CVCC); otherwise, the heaviest syllable receives stress; in case of a tie, the rightmost non-final heavy syllable attracts stress (Hayes 1995, cited by Broselow *et al*. 1997, 49). Following the assumption that syllable weight is determined by moraic structure (Hyman 1985; Hayes 1989), Broselow *et al*. (1997) argued that weightless coda consonants as in Malayalam do not have their own mora but share a mora with the preceding vowel,[29] while weight-bearing coda consonants as in Hindi occupy a head of their own mora, as represented in (35):

(35) Syllable structures for CVC syllables in two languages (based on Broselow *et al*. 1997)

Therefore, the weight of CVC syllables is dependent on coda moraicity. In terms of the moraicity of coda consonants, we can assume the following constraints *C_μ and Weight-By-Position:[30]

(36) Constraints relevant to coda moraicity
*C_μ : Coda consonants are not moraic (Broselow *et al*. 1997)
Weight-By-Position (WBP): Coda consonants are moraic (Hayes 1989)

If *C_μ outranks WBP, coda consonants are not moraic.[31] Recall that CVC syllables do not categorically attract accent in NKK native lexical patterns. Thus, coda consonants do not

---

[29] An alternative representation for a weightless coda is one in which weightless codas attach directly to the syllable node (e.g. McCarthy & Prince 1986; Davis 1994), but Broselow *et al.* (1997) provide empirical evidence for a shared mora by a weightless coda: in Malayalam, vowel duration (both long and short) was significantly shorter in closed syllables than in open syllables, whereas no effect of coda consonants was found on vowel duration in Hindi, which is assumed to have a coda consonant bearing its own mora.

[30] Broselow *et al.* (1997) proposed a mora constraint set, which gives rise to the different moraic structures in different languages: MoraicCoda: All coda consonants must be dominated by a mora; NoSharedMora: Moras should be linked to single segments; and NoCMora: The head of a mora must be a vowel. Depending on the ranking of these constraints, coda moraicity and moraic structure are determined. Since internal moraic structure of CVC syllables in NKK will not be discussed in depth, this chapter uses WBP rather than the mora constraint set for the sake of simpler analyses. I leave the internal moraic structure in NKK for future studies.

necessarily contribute to syllable weight in NKK lexical accentuation, so it seems reasonable to assume that $*C_\mu$ is ranked higher than WBP in NKK.

However, the ranking between the two constraints is not crucial when the faithfulness constraint IDENT (T) outranks the constraints $*C_\mu$ and WBP in native lexical accentuation of existing words. The definition of IDENT(T) is repeated from (3) in Chapter 2.

(37) IDENT (T) (de Lacy 2002):
If mora *x* bears tone T in the input, then the output correspondent of *x* bears T.

Since accented syllables are always realized with a high tone on the surface form, I assume that underlying accent is specified with a high tone. According to the definition in (37), the constraint IDENT (T) forces an underlying high tone (underlying accent) to be realized faithfully on the surface form. Even in cases when WBP outranks $*C_\mu$, heavy CVC syllables are to lose a mora to satisfy IDENT(T) and WTP at the expense of violations of WBP, as illustrated in (38d). Hence, CVC syllables do not necessarily attract a high tone when IDENT (T) dominates WBP. The following tableau illustrates non-moraic coda consonants in lexical accentuation, which results from the ranking of the constraints: WTP, IDENT(T) >> WBP >> $*C_\mu$.

(38) Non-moraic coda in lexical accentuation: WTP, IDENT(T) >> WBP >> $*C_\mu$

Input: |cin.tal.lé.|	WTP	IDENT(T)	WBP	$*C_\mu$
a. cín$_\mu$.tal$_\mu$.le.	*!	*		**
b. cín$_\mu$.tál$_\mu$.le.		*!		**
c. cin$_\mu$tal$_\mu$lé.	*!*			**
☞ d. cin.tal.lé.			**	

The unfaithful candidates (38a) and (38b) are ruled out due to their violations of IDENT(T) and/or WTP. The faithful candidate (38c) also loses because moraic codas incur multiple violations of WTP. Therefore, the candidate (38d) with non-moraic codas, which satisfies WTP and IDENT(T) at the expense of violating WBP, is the winner. The opposite ranking between WBP and $*C_\mu$ also works for the same output in (38).

However, the ranking of WBP >> $*C_\mu$ is crucial in explaining the tendencies toward accented CVC syllables in novel words and existing words, which suggest that coda consonants should add a mora. The following tableau illustrates a case where WBP plays a crucial role in selecting an optimal output.

[31]Gordon (1999; 2004) claims that cross-linguistic variation in weight distinction originates from distinct phonetic properties of rimes in the languages: For example, in Khalka, where CVV syllables are treated as heavy but not CVC syllables, the phonetic distinction in total energy was most visible between VV and other rhymes, while in Finnish, where both CVV and CVC syllables are treated as heavy, the phonetic distinction between VV/VC and other rimes was most visible. Given these facts, Gordon suggested that languages adopt weight distinctions which are most effective phonetically.

(39) Moraic coda in novel words

Input: |tap.kan.pa.|	WTP	IDENT(T)	WBP	*C_μ
a.tap.kán.pa.			*!*	
b tap.kán$_\mu$.pa.			*!	*
c.tap$_\mu$.kán$_\mu$.pa.	*!			
☞d. táp$_\mu$.kán$_\mu$.pa.				**

The candidates (39a) and (39b), which contain non-moraic coda consonants, fail due to violations of WBP. The candidate (39c) with penultimate accent, where non-moraic coda consonants are absent, is also ruled out by WTP. Therefore, double accent is the best choice for /CVC.CVC.CV/ words, satisfying both WTP and WBP.

In the cases where an initial syllable is the only CVC syllable; e.g. *tap.ka.pa.*, the ranking of WBP >> *C_μ is also important to predict the preference for non-penultimate accent. However, evaluation with the constraints in (40) would predict the wrong output: initial accent (40c) rather than double accent (40d). In fact, double accent was preferred to initial accent for the CVC.CV.CV novel forms, although the initial accent can occur in lexical forms of this type. As shown in (40), the candidate with initial accent (40c) satisfies WTP and incurs fewer violations of *H than the candidate with double accent (40d). Therefore, the current system (40) would choose an undesirable output (40c), which is marked with the symbol 👎 in (40). The desirable output (marked with ☹) cannot be chosen due to multiple violations of *H.

(40) |CVC.CV.CV|: double accent vs. initial accent

Input: |tap.ka.pa.|	WTP	WBP	*C_μ	*H
a. tap.ká.pa.		*!		*
b. tap$_\mu$.ká.pa.	*!		*	*
👎 c táp$_\mu$.ka.pa.			*	*
☹ d. táp$_\mu$.ká.pa.			*	*!*

However, if the system includes the constraint ALIGN- R, which dominates *H, the evaluation predicts the correct output, as shown in (41).

(41) Preference for double accent to initial accent

Input: |tap.ka.pa.|	IDENT(T)	WTP	WBP	*C_μ	ALIGN-R	*H
c. táp$_\mu$.ka.pa.				*	*!*	*
☞ d. táp$_\mu$.ká.pa.				*	*	**

The double accented candidate (41d) wins over the initial accented candidate (41c) due to a less serious violation of ALIGN-R.

Returning to the question of what motivates different behaviors of CVV syllables *vs.* CVC syllables in NKK accent placement: CVV syllables always attract accent, while CVC syllables only show a tendency to attract accent, I propose that this difference derives from the ranking status of the relevant constraints. The constraint WTP strictly dominates other relevant constraints, ensuring that CVV syllables always attract accent. However, the current system with rigid constraint ranking cannot provide an account of the probabilistic tendencies: CVC syllable frequently attract accent, but do not always do so. I argue that the incomplete evidence in the lexical accent patterns prevent learners from assuming a strict domination relationship between the two constraints WBP and *C_μ, which determine coda moraicity. These constraints are variably ranked. The following sections will suggest a grammar adopting variable ranking and discuss more details on the variable ranking.

6.3.3 Variation

The system illustrated so far overlooks accent variation in novel forms. The structure-sensitive accent patterns emerging in the novel forms cannot be predicted categorically; e.g., final accent was most frequent in CV.CV.CVC novel forms, but non-final accent patterns were also possible.

Variation has been analyzed using variable constraint ranking as in Anttila (1997): there are certain constraint-pairs for which ranking between the two is free. For example, the variations between penultimate accent and final accent in CV.CV.CVC forms and between penultimate accent and double accent in CVC.CVC.CV forms could be explained by variable ranking of the constraint WBP with respect to other relevant constraints, e.g. *C_μ. Depending on the ranking of these constraints at a particular speech event, coda consonants may or may not contribute to syllable weight, and accordingly either penultimate accent or non-penultimate accent will be assigned. Therefore, variant accent patterns (penultimate vs. non-penultimate) can be produced in accordance with variable constraint ranking, as illustrated in the following tableaux for the cases of CVC.CVC.CV and CV.CV.CVC novel forms.

(42) Penultimate accent vs. Double accent in |CVC.CVC.CV|

i) Non-moraic codas induce penultimate accent: *C_μ >> WBP

Input: |tap.kan.pa.|	WTP	*C_μ	WBP	*H
☞ a. tap.kán.pa.			**	*
b. táp.kán.pa.			**	**!
c. tap$_\mu$.kán$_\mu$.pa.	*!	**		*
d. táp$_\mu$.kán$_\mu$.pa.		*!*		**

ii) Moraic codas induce double accent: WBP>>*C_μ

Input: |tap.kan.pa.|	WTP	WBP	*C_μ	*H
a.tap.kán.pa.		*!*		*
b táp.kán.pa.		*!*		**
c.tap$_\mu$.kán$_\mu$.pa.	*!		**	*
☞ d. táp$_\mu$.kán$_\mu$.pa.			**	**

When *C_μ outranks WBP, as shown in (42i), the candidate with penultimate accent (42ia) will become a winner despite violation of WBP. The other candidates (42ib), (42ic), and (42id) are ruled out due to crucial violations of *H, WTP and *C_μ, respectively. On the other hand, when WBP outranks *C_μ, as in (42ii), the double accented candidate (42iid) in which all codas are moraic will be chosen over the other candidates, since the higher ranked constraints WBP and

WTP rule out the candidates with non-moraic coda consonants (42iia) and (42iib), and the candidate containing an unaccented heavy syllable (42iic), respectively.

On the other hand, we found that final accent, as well as penultimate accent, was common in CV.CV.CVC forms. As in the CVC.CVC.CV novel forms, the variable patterns in CV.CV.CVC forms can also be analyzed using variable ranking of the constraint WBP with respect to other constraints such as NONFINALITY and $*C_\mu$. As shown in (43), penultimate accent is the best choice for CV.CV.CVC forms under the ranking NONFINALITY, $*C_\mu$ >> WBP, while the opposite ranking selects final accent.

(43) Penultimate accent vs. Final accent in |CV.CV.CVC|

i) Non-moraic codas induce penultimate accent: NONFINALITY, $*C_\mu$ >> WBP

Input: |ta.ka.pak.|	WTP	NONFINALITY	$*C_\mu$	WBP	ALIGN-R
a. ta.ká.pak$_\mu$.	*!		*		*
b. ta.ka.pák$_\mu$.		*!	*		
c. ta.ka.pák.		*!		*	
☞d. ta.ká.pak.				*	*

ii) Moraic codas induce final accent: WBP>> NONFINALITY, $*C_\mu$

Input: |ta.ka.pak.|	WTP	WBP	NONFINALITY	$*C_\mu$	ALIGN-R
a. ta.ká.pak$_\mu$.	*!			*	*
☞b. ta.ka.pák$_\mu$.			*	*	
c. ta.ka.pák.		*!	*		
d. ta.ká.pak.		*!			*

When WBP is ranked lower than NONFINALITY and $*C_\mu$, as shown in (43i), the candidate with penultimate accent (43id) will be chosen, despite a violation of WBP, because the higher ranked constraints WTP and NONFINALITY rule out the other candidates (43ia), (43ib), and (43ic). On the contrary, when WBP outranks NONFINALITY and $*C_\mu$, as in (43ii), the final accented candidate with a moraic coda (43iib) wins over the other candidates, because the higher ranked constraints WTP and WBP rule out the candidate containing an unaccented heavy syllable (43iia) and the candidates with a non-moraic coda consonant (43iic) and (43iid).

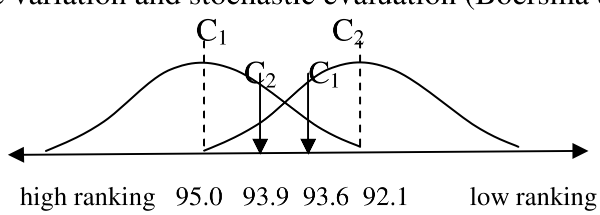
Section 6.3 outlined the constraints which are relevant to the NKK accent alignment in novel words. Most of the accent markedness constraints are assumed to be ranked lower than the faithfulness constraint IDENT(T), which requires underlying accent of the lexical accent patterns to surface faithfully. I argue that the markedness constraints, which do not influence the lexically determined accent patterns, become active in assigning accent in novel words (e.g. Zuraw 2000; Hayes & Londe 2006). Furthermore, the structure-sensitive accent patterns in novel words are not categorical and exhibit variation, which can be explained by variable constraint ranking, as illustrated in this section. Extending Anttila(1997)'s approach to variation, Boersma (1997) and Boersma and Hayes (2001) proposed stochastic ranking: each constraint is provided with a probability distribution on a ranking scale. Under this approach, the variable constraint ranking is attributed to a different probability of different ranking. I assume that stochastic ranking is well suited to analyzing probabilistic distribution of variant NKK accent patterns in novel words.[32] The following section will show how stochastic ranking works for NKK accent patterns.

6.4 Stochastic Ranking

NKK accent patterns in novel words revealed regular tendencies in that CVC syllables tended to carry accent; otherwise, penultimate syllables were the preferred position to be accented. However, a categorical grammar cannot explain these statistical tendencies emerging in novel words. I argue that a stochastic grammar provides an account of the probabilistic accent patterns in novel words. Furthermore, Chapter 4 showed that the structure-sensitive patterning in novel words was consistent with the lexical statistical patterning. This suggests that NKK speakers internalize the lexical statistical patterning and project the patterning onto novel forms. Therefore, I argue that NKK speakers construct a stochastic grammar which is equipped with probabilistic constraint ranking on the basis of the native lexical accent patterning.

In order to model the NKK speakers' knowledge of the probabilistic structure-sensitive accent patterning, I adopt Stochastic OT, as proposed in Boersma 1998; Boersma and Hayes 2001. In accordance with stochastic OT, I assume that the rankings of the constraints are not fixed but rather determined on a continuous scale. Stochastic OT is able to capture statistical patterning and free variation by assuming that the constraints are ranked on a continuous scale, and that candidate evaluation is stochastic with a little bit of noise. This property would generate a situation where constraints overlap when they are close enough each other. When the constraints overlap, free variation occurs via variable ranking, as schematized in (44):

(44) Free variation and stochastic evaluation (Boersma & Hayes 2001)



Including a little random noise in each evaluation of a form results in cases when the two overlapping constraints swap places in the hierarchy for a given speech event. Constraint C_1 has a ranking of 95.0, outranking constraint C_2 which has a ranking value of 92.1 (these numbers are arbitrary). However, in some cases, e.g. when C_2 gets a ranking of 93.9 and C_1 gets a ranking of 93.6, C_2 outranks C_1. Greater overlap in the ranges of the constraints indicates a greater possibility that the constraints will swap places in a specific evaluation. Therefore, this system can implement the probabilistic tendency for C1 to outrank C2 rather than completely free ranking between the two. A specific example relevant to NKK accentuation is as follows: suppose that the ranking values of the constraints NONFINALITY and ALIGN-R are 98.5 and 94.6, respectively. If the probability for NONFINALITY which outranks ALIGN–R is 86.5%, the opposite ranking of the constraints could occur on some occasions (13.5%), where the two constraints overlap. This indicates that penultimate accent would win over final accent 86.5% of the time,

[32] Other competing models have been proposed for variable ranking: e.g. Harmonic Grammar (Potts, Pater, Jesney, Bhatt & Becker 2010); Maximum Entropy (Hayes 2008).

but final accent should appear 13.5% of the time in cases when a word does not contain heavy syllables.

In the subsequent sections, I will show how the structure-sensitive tendencies in NKK accent patterns are modeled using stochastic ranking. In order to obtain stochastic rankings for the NKK accent patterns, I used the Gradual Learning Algorithm[33] (GLA)(Boersma1997; Boersma & Hayes 2001), which has been argued to successfully handle statistical tendencies and noisy learning data, based on stochastic OT (Zuraw 2000; Boersma & Hayes 2001; Hayes and Londe 2006; Apoussidou 2007).

6.4.1 The grammar of NKK accent patterns in novel words

This section will provide a theoretical model of probabilistic NKK accent patterns emerging in novel words, using Boersma's GLA. The GLA was originally designed to learn a stochastic grammar from variable data. The algorithm is error-driven so learning occurs when an output is incorrect. The constraint violations of incorrect winners are compared to those of the correct output and constraint rankings are adjusted accordingly. Given the finding that the GLA works fairly well in learning a grammar of statistical patterns (e.g. Zuraw 2000; Boersma & Hayes 2001; Hayes and Londe 2006; and Apoussidou 2007), if the GLA successfully learns an NKK accent grammar based on the NKK novel learning data, I assume that the resulting ranking will represent the grammar of NKK speakers.

All the constraints discussed in 6.3 are initially ranked at 100. The training data set consists of 40 pairs of given underlying forms together with surface forms, as listed in (45): four different underlying forms of trisyllabic words consisting of syllable structure combinations CV.CV.CV, CVC.CV.CV, CV.CV.CVC, and CVC.CVC.CV. For each underlying form, five different accent patterns in corresponding furface forms are logically possible: no accent, initial accent, penultimate accent, final accent, and double accent. Surface forms include moraic codas and non-moraic codas. The input frequencies were provided based on the results in Experiment 1.

(45) Training data

<table>
<tr><td>underlying forms</td><td>|cv.cv.cv|</td><td>|cvc.cv.cv|</td><td>|cv.cv.cvc|</td><td>|cvc.cvc.cv|</td></tr>
<tr><td>surface forms</td><td>/cv́.cv.cv/</td><td>/cv́c.cv.cv/</td><td>/cv́.cv.cvc/</td><td>/cvć.cvc.cv/ /cv́c.cvC.cv/</td></tr>
<tr><td></td><td>/cv.cv́.cv/</td><td>/cvc.cv́.cv/</td><td>/cv.cv́.cvc/</td><td>/cvc.cv́c.cv/ /cvc.cv́C.cv/</td></tr>
<tr><td></td><td>/cv.cv.cv́/</td><td>/cvc.cv.cv́/</td><td>/cv.cv.cv́c/</td><td>/cvc.cvc.cv́/ /cvc.cvC.cv́/</td></tr>
<tr><td></td><td>/cv́.cv́.cv/</td><td>/cv́c.cv́.cv/</td><td>/cv́.cv́.cvc/</td><td>/cv́c.cv́c.cv/ /cv́c.cv́C.cv/</td></tr>
<tr><td></td><td>/cv.cv.cv/</td><td>/cv́C*.cv.cv/</td><td>/cv́.cv.cvC/</td><td>/cv́C.cvc.cv/ /cv́C.cvC.cv/</td></tr>
<tr><td></td><td></td><td>/cvC.cv́.cv/</td><td>/cv.cv́.cvC/</td><td>/cvC.cv́c.cv/ /cvC.cv́C.cv/</td></tr>
<tr><td></td><td></td><td>/cvC.cv.cv́/</td><td>/cv.cv.cv́C/</td><td>/cvC.cvc.cv́/ /cvC.cvC.cv́/</td></tr>
<tr><td></td><td></td><td>/cv́C.cv́.cv/</td><td>/cv́.cv́.cvC/</td><td>/cv́C.cv́c.cv/ /cv́C.cv́C.cv/</td></tr>
<tr><td></td><td></td><td>/cvc.cv.cv/</td><td>/cv.cv.cvc/</td><td>/cvc.cvc.cv/</td></tr>
</table>

* capital C represents a moraic coda (c_μ).

[33] Boersma's Gradual Learning Algorithm is available in OTSoft 2.3.1 at http://www.linguistics.ucla.edu/people/hayes/otsoft/.

There were a total of 100,000 learning trials. Learning was done with an evaluation noise of 2.0 and an initial plasticity of 2.0 decreasing at each stage and ending up at 0.001.[34] I ran the algorithm for multiple trials on the same data and all trials yielded similar results. In a representative run, the algorithm obtained the following ranking values for the constraints:

(46) Ranking values generated by the GLA

Constraints	Ranking values
WEIGHT-TO-TONE PRINCIPLE (WTP)	110.000
$H]_{PW}$	106.159
WEIGHT-BY-POSITION (WBP)	100.221
NONFINALITY	100.058
$*C_{\mu}$	99.779
ALIGN-R	95.783
*H	-5,763.625

A considerable distance between two constraints suggests strict domination relations between the two constraints. For example, the ranking values of WTP and NONFINALITY are 110 vs. 100, which implies that WTP strictly dominates NONFINALITY. This resulting grammar shows that the algorithm successfully learned the proposed constraint ranking NONFINALITY >> ALIGN-R >> *H, which is argued to derive the predominant penultimate accent in words without heavy syllables. The ranking values of these constraints are considerably far apart: 100 vs. 95 vs. -5,763.

On the other hand, ranking values close to each other or nearly identical indicate that free ranking of the two constraints is highly possible. The ranking values of WBP, NONFINALITY, and $*C_{\mu}$ in (46) are very similar: all of them are near 100 and the distances between the two constraints in three pairs (WBP vs. NONFINALITY; NONFINALITY vs. $*C_{\mu}$; $*C_{\mu}$ vs. WBP) are all less than ·5. In the previous section, it was argued that the variable ranking of these constraints cause variation in the accent patterns of novel forms: penultimate accent vs. non-penultimate accent. Thus, the close ranking values of these three constraints also suggest that the algorithm generalized the grammar appropriately.

[34] The schedules that I employed for plasticity are as follows:

Stage	Trials	Plasticity	Noise
1	25000	2	2
2	25000	0.2	2
3	25000	0.01	2
4	25000	0.001	2

In addition, the algorithm computed the probabilistic rankings of these three constraints, as shown in (47), which helps us to envision the output frequencies of variant accent patterns.[35]

(47)

$$\begin{tikzcd}
\text{WBP} \arrow[d, "\cdot 524"] \arrow[dd, bend left=60, "\cdot 563"] \\
\text{N}\textsc{onfinality} \arrow[d, "\cdot 539"] \\
{*}\text{C}\mu
\end{tikzcd}$$

The three-way pairwise ranking probabilities are all near 50%, which means that chances of ranking alternation would be about 50%. In other words, variation is very likely to occur. For the sake of simple illustration of what the ranking probabilities imply in the grammar, I hypothesize that only a two-way interaction between the two constraints, WBP and NONFINALITY, exists. In this case, the ·524 probability ranking of WBP over NONFINALITY indicates that there is a probability of ·524 that the grammar will generate output /ta.ka.pá$\mathbf{k}_{\mu}$/, given an input like |.ta.ka.pak|, as shown in (48). The opposite ranking will produce an output /ta.ká.pak/ with a probability ·476. The following tableau illustrates how the evaluation with stochastic rankings generates these variations.

[35] The pairwise ranking probabilities are given below. In the table, the numbers indicate the probability at which the constraint in the row headings outranks the constraint in the column headings.

	..H..]$_{PW}$	WBP	NONFINAL	*Cμ	ALIGN-R	*H
WTP	0.914	0.999	0.999	0.999	>.999	>.999
..H..]$_{PW}$		0.983	0.985	0.989	0.999	>.999
WBP			0.524	0.563	0.942	>.999
NONFINAL				0.539	0.935	>.999
*Cμ					0.922	>.999
ALIGN-R						>.999

(48) Ranking probabilities and variation: Case I

.476

.524

Input: |.ta.ka.pak|	WTP	..H..]$_{PW}$	WBP	NonFinal	Align-R	*H
☞.ta.ka.pá$\mathbf{k}_\mu$ (.524)				*(!)		*
☞.ta.ká.pak (.476)			*!		*	*
.ta.ka.pák			*!	*(!)		*
.ta.ká.pa$\mathbf{k}_\mu$	*!				*	*

When the computation included the three-way interaction of all three constraints in (47), WBP, NONFINALITY, and *C_μ, the algorithm predicted the occurrence of /ta.ka.pá$\mathbf{k}_\mu$/ at 38% of the time and the occurrence of /ta.ká.pak/ at 59% of the time.

The variable rankings between the constraints WBP and *C_μ also give an account of the accent variation in words beginning with CVC syllables: penultimate accent vs. double accent. The ·563 probability ranking of WBP over *C_μ indicates that there is a probability of ·563 that WBP will dominate *C_μ, and that the grammar will generate the output /tá$\mathbf{p}_\mu$.ká.pa/ at a probability of ·563, given an input like |.tap.ka.pa|, as illustrated in the following tableau.

(49) Ranking probabilities and variation: Case II

·437

·563

Input: |.tap.ka.pa|	WTP	..H..]$_{PW}$	WBP	NonFinal	*Cμ	Align-R	*H
☞.tá$\mathbf{p}_\mu$.ká.pa (0.563)					*(!)	*	**
☞.tap.ká.pa (0.437)			*!			*	*
.táp.ká.pa			*!			*	**(!)
.ta$\mathbf{p}_\mu$.ká.pa	*!				*	*	*

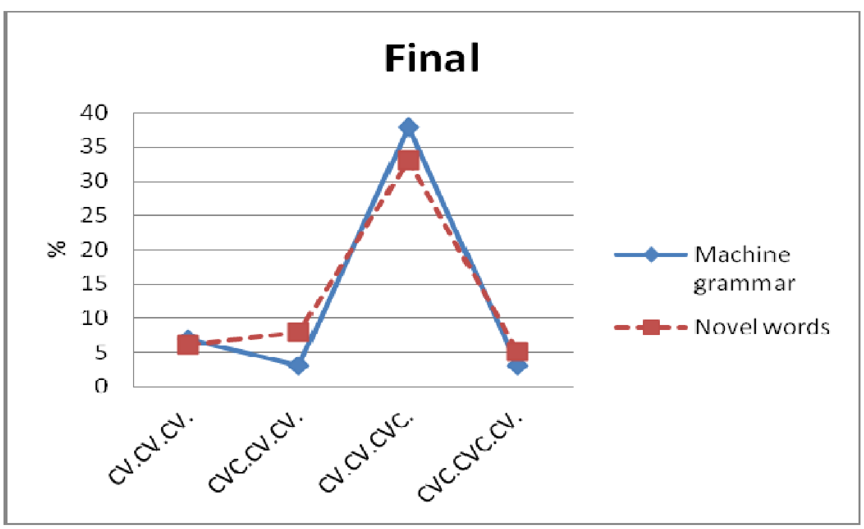
After each run, every underlying form was submitted to the resulting grammar 1,000,000 times to produce output probabilities. In order to assess whether the grammar generated by the algorithm fits the data, the output probabilities predicted by the grammar are compared to the input frequencies of the novel word data. The comparison of the two groups is presented in (50).

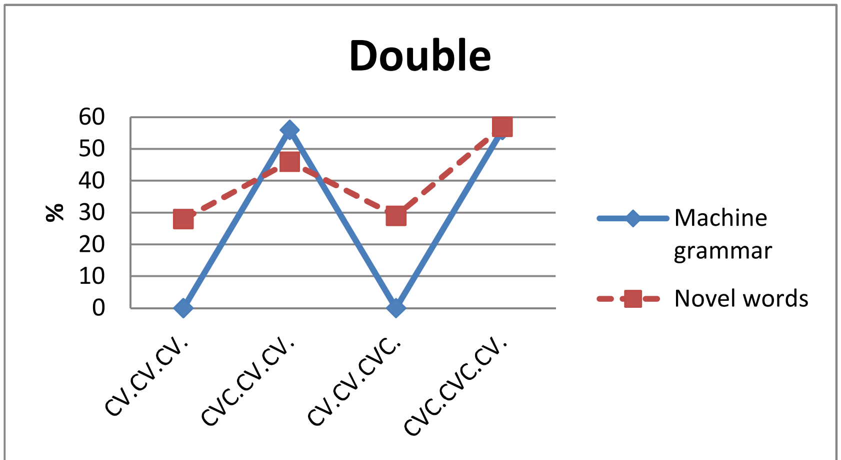
(50) Input frequencies (Novel words) vs. Output frequencies (by Machine-ranked grammar)

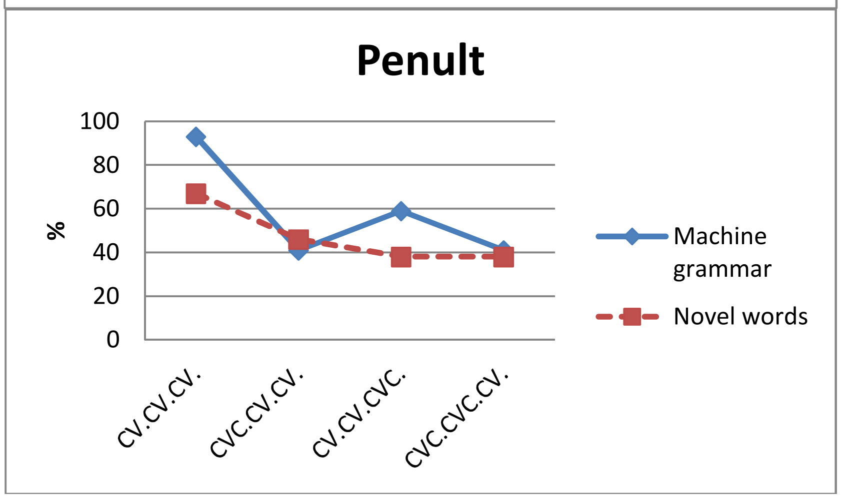
	Accent type	Input (%)	Output (%)
CV.CV.CV.	penult	66.7	93.4
	double	27.5	0
	final	5.8	6.6
CVC.CV.CV.	penult	46.2	40.9
	double	46.2	55.8
	final	7.6	3.3
CV.CV.CVC.	penult	37.5	59.3
	double	29.2	0
	final	33.3	40.7
CVC.CVC.CV.	penult	38.1	40.9
	double	56.8	55.8
	final	5.1	3.3

Although the exact quantitative match with the novel word data was not obtained by the machine grammar, the grammar correctly achieved the overall proportional generalizations: double accent and final accent were more likely in words with word-initial CVC syllables and with word-final CVC syllables, respectively; otherwise, penultimate accent was most common. The following figures illustrate graphically a fairly good match of the output frequencies to the novel word data:

Figure 43 Matchup of the resulting grammar to the novel word data







With regard to final accent, the machine grammar made a fairly good match to the frequencies in the input novel words. Nevertheless, the grammar failed to predict the occurrence of double accent in the novel forms CV.CV.CV and CV.CV.CVC. In contrast, the grammar predicted more frequent penultimate accent in the corresponding word types than in the actual input novel forms. This mismatch should be attributable to the presence of phonetically-driven variations in the novel words. Chapter 3 showed that words beginning with a high vowel were more likely to carry double accent than words beginning with a low vowel. The following chapter will also show that phonetic quality serves as another factor influencing variation in NKK accent patterns: double accent is more likely in words beginning with aspirated consonants. Hence, I speculate that the machine grammar predicted double accent with a lower frequency due to the absence of constraints relevant to the phonetic effects. If the grammar includes the relevant constraints, it will achieve a fairly good match with the input frequencies. Therefore, I conclude that the resulting grammar by the GLA provides a good fit with the NKK accent patterns in novel words.

6.4.2 Learning the grammar

The previous section proposed a grammar with stochastic rankings which produces the NKK syllable structure-sensitive accent tendencies in novel words. Chapter 4 demonstrated that the accent structure-sensitivity is actually reflected in the statistical accent patterns in existing words, which implies that learning of the grammar should be possible via exposure to the lexicon. This section will test whether the suggested grammar is indeed learnable based on the lexical accent patterns in existing native words. Thus, on this trial I provided the GLA with the input frequencies from the corpus data (reported on in Chapter 4) as learning data, as follows in (51). I assumed that the corpus data is representative of the data from the NKK lexicon to which learners are exposed.

(51) Learning data based on the corpus: Input frequencies[36]

word type	frequency			
	penult	double	final	initial
CV.CV.CV.	138	20	4	6
CVC.CV.CV.	78	12	0	1
CV.CV.CVC.	4	6	13	2
CVC.CVC.CV.	36	16	4	1

If the suggested grammar of the accent patterns in novel words is learnable based on the lexical patterns, the GLA with the corpus data will generate a grammar similar to the grammar which we have achieved based on the novel word data.

On the other hand, existing words in the corpus are assumed to be different from novel words in that underlying accent is present in existing words. Including lexical specifications for accent in the underlying form, the training data set consists of 200 pairs of given underlying forms together with surface forms (40 pairs (as listed in (45) in the section 6.4.1) with 5 different underlying accents: no accent; initial accent; penultimate accent; final accent; and double accent). Furthermore, a faithfulness constraint IDENT(T), which forces an underlying accent to surface faithfully, should be relevant to accentuation in existing words. Thus, the input file included the constraint in the simulations. As for the underlying forms which have no accent information, the frequencies of each accent pattern in (51) were provided for each word type in order to let the GLA generalize the patterning of unaccented words based on the lexical frequency. I ran the GLA for several trials and all trials yielded similar results. A representative run obtained the following ranking values in (52).

[36] Only type frequencies are used.

(52) Learned grammar: Ranking values generated by the GLA

Constraints	Ranking values
IDENT(T)	118.000
H]$_{PW}$	116.000
WEIGHT-TO-TONE PRINCIPLE (WTP)	112.000
*C_μ	100.313
WEIGHT-BY-POSITION (WBP)	99.687
NONFINALITY	80.935
ALIGN-R	75.898
*H	73.393

The resulting grammar shows that all constraints except for *C_μ and WBP maintain a distance of at least two numeric values between the adjacent constraints, which means that the constraints would be in a relationship of strict domination. Given the lexical frequency, the GLA successfully generalized the ranking NONFINALITY >> ALIGN-R >> *H, which is crucial for penultimate accent assignment in words without heavy syllables. The ranking values of the two constraints *C_μ and WBP are very close, less than one value apart, which was also similar to the target grammar, obtained in the previous section. The converted ranking probability of this pair is ·588, which means that the constraint *C_μ would outrank WBP in 59% of speech events.[37] This implies that coda consonants would not count as moraic, and therefore CVC syllables will not attract accent, 59% of the time. The probability of the opposite ranking WBP>> *C_μ is ·412, so CVC syllables will attract accent at 41% of the time because CVC syllables are counted as heavy under this ranking.

In comparing the grammar obtained with the existing forms and the target grammar obtained with novel word data, the grammar learned on the basis of the corpus data shows a fairly good match with the target grammar in general, as shown in (53):

[37] The computed pairwise ranking probabilities are as follows:

Ranking	Probability
WTP >> *C_μ	>.999
***C_μ >> WBP**	**0.588**
WBP >> NONFINAL	>.999
NONFINAL >> ALIGN-R	0.963
ALIGN-R >> *H	0.813

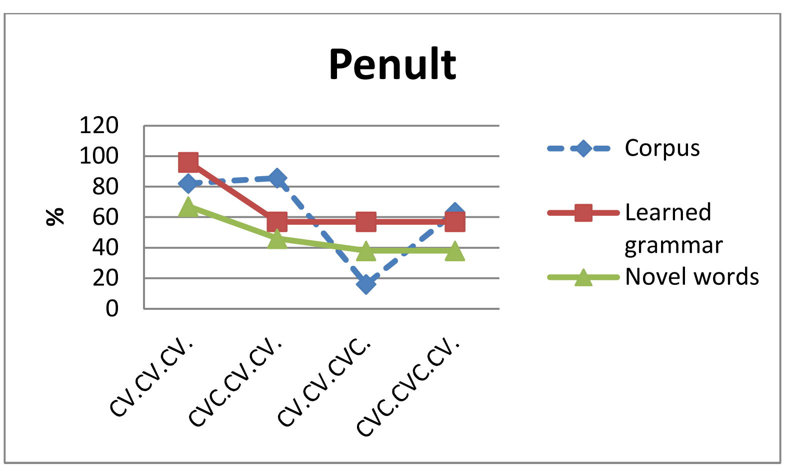
(53) Target grammar (novel words) vs. Learned grammar (the corpus)

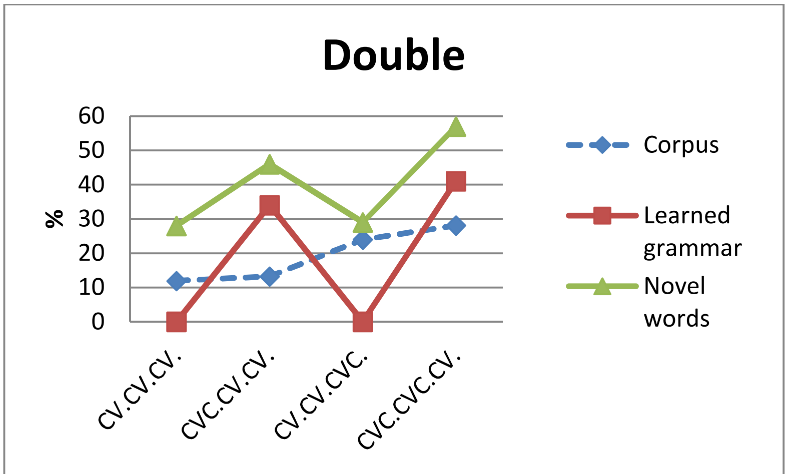
Target grammar		Learned grammar	
Constraints	Ranking values	Constraints	Ranking values
WTP	110.000	$H]_{PW}$	116.000
$H]_{PW}$	106.159	WTP	112.000
WBP	100.221	$*C_\mu$	100.313
NONFINALITY	100.058	WBP	99.687
$*C_\mu$	99.779	NONFINALITY	80.935
ALIGN-R	95.783	ALIGN-R	75.898
*H	-5,763.625	*H	73.393

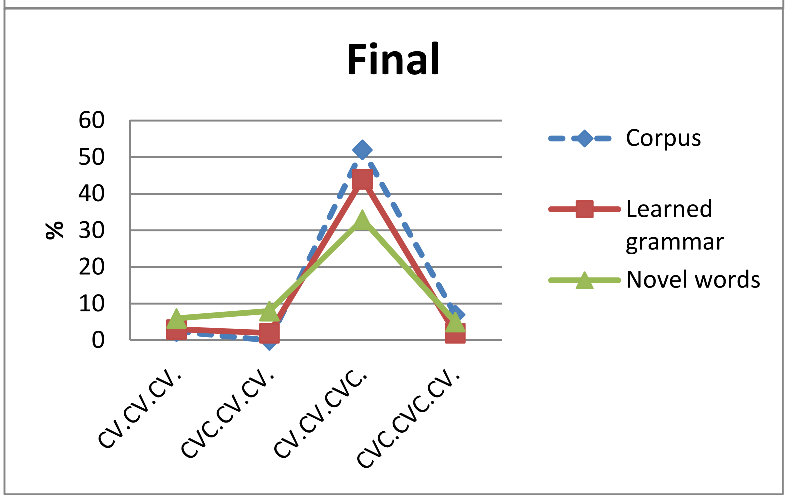
The constraints WTP and $H]_{PW}$ are undominated in both grammars but the hierarchical order of these two constraints is switched in the learned grammar. However, the relative ranking between the two constraints is not important, since alternating ranking does not bring about any change in the accent patterns. The learned grammar maintains the crucial hierarchical order in the rankings of the remaining constraints, as in the target grammar, except for the constraint $*C_\mu$. CVC syllables were less likely to carry accent in the existing data than in the novel data; due to the top-ranked constraints IDENT (T) and WTP, CVC syllables without underlying accent must be realized with a low tone (unaccented). As the virtual learners encounter more unaccented CVC syllables, the constraint $*C_\mu$ is promoted gradually. Nonetheless, the ranking values of the two constraints WBP and $*C_\mu$ are very close, which means that the switched ranking is also probable. Thus, the resulting accent patterns would not be so different from the patterns in the target grammar, whose ranking values between the two constraints were also near each other. Therefore, the difference in the ranking values of $*C_\mu$ between the two grammars should not be meaningful in terms of the resulting accent patterns.

The following figures illustrate that the accent distributions predicted by the learned grammar are similar to those in novel words, which suggests that learning of the proposed grammar is indeed possible via the lexicon. In the figures, the corpus input frequencies are compared with the output frequencies predicted by the learned grammar.

Figure 44 Matchup frequency of the learned grammar to the novel words







Final accent shows a fairly good match among all three groups. In the distribution of penultimate accent and of double accent, the learned grammar also produced similar patterns to the patterns in the novel words, although the grammar was not able to mimic the numeric quantities. The discrepancy arose mainly from more frequent double accent in novel words than the grammar predicted, which can be traced to phonetically-driven variation in accent patterns, as discussed earlier: double accent was more likely in a word beginning with an element with inherently high pitch, such as aspirated onset consonants or high vowels. As suggested in the previous section, if constraints which give rise to the phonetic effects on accent patterns, were included in the simulation, they should compensate for the mismatch of double accent assignment. Moreover, the learned grammar predicted output frequencies similar to the patterns in novel words rather than those in the input corpus data, although the simulation was run based on the corpus data. All these together strongly suggest that it is possible to generalize the structure-sensitive accent patterns on the basis of the native lexical patterns, though considerable exceptions are present. Therefore, the target stochastic grammar, which is argued to determine the accent patterns in novel words, is indeed learnable via the native lexicon.

6.5 Accenting Existing Words vs. Novel Words

The previous section suggested that the native NKK lexicon provides sufficient evidence for the grammar used to assign accent to the novel words. I argue that most of the markedness constraints responsible for the structure-sensitive tendencies are ranked lower than the faithfulness constraint IDENT(T). Hence, the constraints remain inactive when accenting existing words, but the constraints become active when it comes to new words, assuming that the new words lack underlying accent. This section will describe how accent is assigned on existing words *vs.* new words under the proposed stochastic grammar.

In terms of accent placement in existing words, NKK learners can assume that the accent position on the surface form faithfully reflects the underlying accent under the ranking IDENT(T) >> accent related markedness constraints. This assumption is in accordance with the Lexicon Optimization proposal of Prince & Smolensky (1993: 191), in that the underlying form of a word is determined by evaluating different possible underlying forms with respect to a surface form which is optimal in the ranking of the language. The optimal surface form is determined by the ranking of structural constraints, and the appropriate underlying form for this surface form is determined by faithfulness: the most faithful underlying-surface pair is the most harmonic one, and chosen as the optimal pair.

Additionally, lexical constraints are necessary to account for accent assignment in a specific existing lexical item. I adopt the lexical constraints which connect *meaning* to underlying forms, as proposed in Boersma (2001) and in Apoussidou (2007). They suggest that these lexical constraints need not be innate but can be language-specific. The lexical constraints are formulated as follows:

(54) Lexical constraints on underlying accent specification[38]

Don't connect the meaning of 'xy' to the form |XY| that is specified/unspecified for accent.

The lexical constraints work against the connection of a meaning to an underlying form that is specified as accented or unaccented. For example, for a lexical item *kámani* 'straw rice-bag', two constraints in (55) are suggested on the relation between a meaning and an underlying form.

(55) Lexical constraints on a lexical item *kámani* 'straw rice-bag'

|kamani| 'straw rice-bag': Don't connect the meaning 'straw rice-bag' to an unaccented stem |kamani|.*

|kámani| 'straw rice-bag': Don't connect the meaning 'straw rice-bag' to an accented stem |kámani|.*

Finally, hearing an output like [kámani] produced by another NKK speaker, NKK learners would choose the most faithful underlying-surface pair and consider the form |kámani|

[38] The constraints are a revised version for accent specification based on Apoussidou (2007:170). Zuraw (2000) uses a constraint USELISTED, which requires that a single lexical entry rather than a string of morphemes be used as input.

as the underlying form. Accordingly, NKK learners will rank *|kamani| 'straw rice-bag' higher than *|kámani| 'straw rice-bag'.[39] An underlying form is reconstructed by evaluation of underlying-surface pair candidates in recognition of an existing lexical item *kámani* 'straw rice-bag', as illustrated in (56).

(56) Recognition of *kámani*

Surface form /.ká.ma.ni/: Input ~Output	IDENT(T) (0.922)	..H..]$_{PW}$ (0.984)	WTP (0.998)	*|kamani| (0.979)	*Cμ (0.577)	WBP (0.979)	*|kámani| (0.996)	NonFinal (0.965)	ALIGN-R (0.814)	*H
☞|.ká.ma.ni| ~ /.ká.ma.ni/ (1)							*		**	*
|.ka.má.ni| ~ /.ká.ma.ni/	*!								**	*
|.ka.ma.ní| ~ /.ká.ma.ni/	*!								**	*
|.ka.ma.ni| ~ /.ká.ma.ni/				*!					**	*

* The number in parentheses after the constraint labels indicates the probability that each constraint outranks the next one down.

The system selects the most faithful pair |.ká.ma.ni| ~ /.ká.ma.ni/ as an optimal output. Therefore, for the surface form /.ká.ma.ni/, the underlying form|.ká.ma.ni| was chosen.

In terms of production, speaking is initiated with the meaning the speaker wants to express. Thus, in the production grammar, candidates must include *meaning* in addition to underlying forms and surface forms. Accent production of the word *kámani* 'straw rice-bag' is generated via meaning-to-form evaluation under the same system as in recognition, whereas surface form candidates are different from candidates in recognition. In recognition, only the surface forms with initial accent are considered as candidates since NKK listeners have already perceived the initial-accented output. However, in production, all possible accents are considered as surface form candidates. The evaluation of meaning-to-form is illustrated in (57) for the production of a word ''straw rice-bag'. The tableau in (57) omits the following constraints H]$_{PW}$; WTP; *Cμ; and WBP, which do not operate crucially in selecting an output from the listed candidates.

[39] NKK learners are in the process of developing a grammar, and thus might produce incorrect forms. If they encounter produced forms mismatched with perceived forms, they will detect an error and adjust their constraint ranking.

(57) Production of 'straw rice-bag'

'straw rice-bag'	IDENT(T) (0.922)	*|kamani| (0.979)	*|kámani | (0.996)	NONFINAL (0.965)	ALIGN-R (0.814)	*H
☞a. 'straw rice-bag'|.ká.ma.ni| /.ká.ma.ni/			*		**	*
b. 'straw rice-bag' |.ká.ma.ni| /.ka.má.ni/	*!		*		*	*
c. 'straw rice-bag' |.ka.ma.ni| /.ka.má.ni/		*!			*	*
d. 'straw rice-bag' |.ka.ma.ni| /.ká.ma.ni/		*!			**	*
e. 'straw rice-bag' |.ka.ma.ni| /.ka.ma.ní/		*!		*		*
f. 'straw rice-bag' |.ka.ma.ni| /.ká.má.ni/		*!			*	**

The lexical constraint *|kamani| 'straw rice-bag' is crucial to rule out all the candidates with the unaccented underlying form |.ka.ma.nil(57c~57f). The constraint IDENT(T) does not allow unfaithful underlying-surface pairs like (57b). Finally, the most faithful surface form to the underlying form (57a) is selected.

On the other hand, if a NKK speaker has never heard this word or she is not sure whether the underlying form is accented or unaccented (e.g. because the word is infrequently used), her grammar would consider the unaccented underlying form as candidates. In other words, the opposite ranking of the lexical constraints: *|kámani| 'straw rice-bag' >> *|kamani| 'straw rice-bag' would be also possible. This scenario is highly plausible because the accent position of an individual NKK lexical item is not fully predictable. Consequently, unfamiliarity or uncertainty could give rise to variation in lexical accent patterns even for existing words. Actually, an alternative pattern, penultimate accent, is possible for the word *kámani* 'straw rice-bag'. Thus, if we hypothesize that NKK learners hear *kamani* with initial accent 50% of the time and with penultimate accent 50% of the time, then the learners would construct a stochastic ranking of the lexical constraints based on the frequency as their learning proceeded. The following tableau envisions the hypothetical stochastic grammar of the variations of the word *kamani*.

(58) Variations in a word ‘straw rice-bag’

·5

·5

‘straw rice-bag’	IDENT(T) (0.922)	*|kámani| (0.5)	*|kamani| (0.5)	NONFINAL (0.965)	ALIGN-R (0.814)	*H
☞ a.‘straw rice-bag’|.ká.ma.ni|/.ká.ma.ni/ (0.5)		*!			**	*
☞ b.‘straw rice-bag’|.ka.ma.ni|/.ka.má.ni/(0.5)			(*!)		*	*
c. ‘straw rice-bag’|.ka.ma.ni| /.ká.ma.ni/			(*!)		**!	*

Once the underlying form of a word is assumed to be unaccented, the constraint IDENT(T) becomes irrelevant. Then, the lexical constraint *|kámani| is considered as being ranked over *|kamani|, which rules out the candidate (58a) containing the accented underlying form. Also, the lower ranked markedness constraints (NONFINALITY; ALIGN-R; and *H) become active in the selection of the optimal output. Hence, the candidate (58b) wins under the evaluation of the lower ranked constraints. Otherwise, the hypothesized probability ranking, 50% of the lexical constraint *|kamani| over *|kámani|, predicts that the candidate (58a) will be chosen alternatively as an optimal output, as the constraint *|kamani| rules out the candidates (58b) and (58c) (indicated by the symbol *! in the parentheses).

In production of novel words, the lexical constraints as well as the faithfulness constraint IDENT(T) become irrelevant. Accordingly, the markedness constraints take over in accent alignment. As illustrated in (59), the accent position of the novel word *takapa* is determined by interaction of the markedness constraints only.

(59) Production of a novel word *takapa*

	Ident(T) (0.922)	*|kamani| (0.979)	|kámani | (0.996)	NonFinal (0.965)	Align-R (0.814)	*H
☞|.ta.ka.pa| ~ /.ta.ká.pa/					*	*
|.ta.ka.pa| ~ /.tá.ka.pa/					**!	*
|.ta.ka.pa| ~ /.ta.ka.pá/				*!		*
|.ta.ka.pa| ~ /.tá.ká.pa/					*	**!

Therefore, the grammar predicts that penultimate accent will be assigned in a word *takapa* most of the time since the candidates with other accent patterns commit a crucial violation of the accent markedness constraints.

To summarize, accent placement of novel words is constrained only by the relevant markedness constraints. In contrast, in the accentuation of existing words, the faithfulness constraint IDENT (T) and lexical constraints play a crucial role in accent placement. Although the constraint IDENT (T) and the lexical constraints obscure the effects of lower-ranked markedness constraints, the markedness constraints and their relative rankings are reflected in accent variation present in existing words, and in the statistical patterning in the lexicon.

6.7 Accent Patterns in Loanwords Revisited

The syllable structure-sensitive tendencies in novel words are actually consistent with the regular patterns emerging in loanwords. The NKK accent patterns of loanwords have been studied by several scholars including Kenstowicz and Sohn (2001) and Chung (2001, 2006), who agreed on the following patterns in loanwords: i) closed syllables generally attract accent (e.g. *kawún* 'gown'; *hóŋkʰóŋ* ‘Hong Kong’; *sɨtʰakʰíŋ* 'stocking'; *kʰélkəri* ‘Calgary’); ii) without closed syllables, penultimate syllables tend to be accented (e.g. *raítʰa* 'lighter'; *apʰátʰɨ* ‘apartment’). When a word contains more than one heavy syllable, however, there is variation (e.g. *kʰéptʰín* ‘captain’; *nepkʰín* ‘napkin’). In addition, a few true exceptions to these patterns exist: e.g. *panamá* ‘banana’; *kʰamerá* ‘camera’; *tʰəksíto* ‘tuxedo’.

It has generally been assumed that the structure-sensitive patterns of loanwords cannot be attributed either to native phonology or to the influence of the source language. Accent placement in loanwords is not necessarily consistent with stress placement in the source language (e.g. *sɨtʰakʰíŋ* vs. *stócking*). Furthermore, coda-weight sensitive accent in loanwords cannot be categorically predicted by NKK native phonology: there is no direct evidence that coda consonants attract accent in native words (e.g. *cintallé* ‘azalea’; *camcári* ‘dragon fly’). For this reason, Chung (2006) provides an OT analysis of the NKK loanword patterns with ‘re-ranked’ constraints and argues that the regular patterning of loanwords represents the emergence of universal grammar (e.g. Kenstowicz and Sohn 2001).

However, the findings in this dissertation suggest that the accent patterns in loanwords should be attributable to the native language association between syllable weight and accent, which is reflected in the lexical statistical patterning. Thus, ‘reranking’ of the constraints would not be necessary in the loanword phonology. The patterns in loanwords can be accounted for by the uniform ranking of the constraints as in the native phonology, as suggested in the section 6.3. Furthermore, although the native accent patterns are determined lexically in NKK, the native accent patterns manifest evidence of universal principles indirectly via lexical frequency. In addition, the syllable structure effects emerging in on-line accent assignment in newly adopted words (novel words) should persist in off-line accent assignment after lexicalization (as in loanwords).

6.8 Conclusion

This chapter provided formal analyses of probabilistic syllable structure-sensitive NKK accent patterns which emerged in novel words, based on stochastic constraint rankings. The stochastic ranking renders a model which captures the statistical (non-categorical) tendencies emerging in novel words.

In terms of learning a phonological grammar, a markedness-over-faithfulness (M>>F) bias has been suggested as the initial state by many researchers (e.g. Gnanadesikan 2004; Hayes

2004; Prince & Tesar 2004). The initial state is postulated to have the ranking of all the relevant markedness constraints (e.g. NONFINALITY, ALIGN-R) over faithfulness constraints (e.g. IDENT(T)). As learners are exposed to exceptional cases, they revise their grammar by promoting a relevant faithfulness constraint, IDENT(T) above the markedness constraints. However, accent patterns in novel words manifest the markedness constraints ranked lower than IDENT (T), which will take over when accenting newly adopted words. The presence of the markedness constraints is also manifested by data from child phonology (Chung 2007) and second language acquisition (H-J Kim 2003). Chung (2007) reported that NKK children aged between 25 and 30 months showed preferences for accented CVC syllables, which indicates that the markedness constraints WTP (high tone on heavy syllables) and WBP (moraic coda) operate actively in child phonology. H-J Kim (2003) also demonstrated that error patterns in NKK learners of English stress were not random but systematic, showing tendencies toward penultimate stress and stressed CVC syllables. This provides additional evidence for the markedness constraints NONFINALITY, ALIGN-R as well as WTP and WBP.

I proposed that the ranking among these markedness constraints is stochastic on the basis of frequency of existing patterns in the lexicon, which suggests that NKK learners also re-rank the markedness constraints gradually in accordance with the lexical tendencies in the existing words. Furthermore, I have demonstrated that the GLA successfully learned the proposed stochastic grammar with the training data from the corpus. This suggests that the stochastic grammar, which is argued to guide accent patterns in novel words, should be learnable via exposure to the lexicon. This also implies that the positive evidence for the markedness constraints in the lexicon should be sufficient to account for the accentuation of novel words. In addition, the structure-sensitive tendencies emerged more strongly in novel words than in the lexical statistical patterning, which reflects learning which is biased toward the structure-sensitive patterns: NKK learners do not simply replicate the lexical frequency but, rather, their accentuation of new words reflects their grammar.

Chapter 7 Phonetic Effects on NKK Accentuation of Novel Words

7.1 Introduction

Chapter 3 showed that NKK speakers' accentuation of newly adopted words is not random, but is partially determined by the syllable structure of the words: heavy syllables tend to be accented. Without heavy syllables, penultimate accent was the preferred pattern in novel words: more than 80% of words which consist of only light syllables were assigned penultimate accent. However, the accent patterning was not categorical: about 20 % of words had non-penultimate accent despite the predominant preference for penultimate accent. Chapter 3 also showed that vowel quality influenced accent patterns in novel words: words beginning with a high vowel were more likely to carry double accent than words beginning with a low vowel. This suggests that phonetic quality (e.g. F0) is a relevant factor at least in deriving variation in the accent patterns.

Chapter 7 investigates phonetic factors that influence NKK accentuation in novel words and presents empirical evidence that aspiration plays a role in the assignment of a pitch accent pattern to novel words. Non-penultimate accent, which was rarely found in a word without aspirated consonants, was chosen when a word contained aspirated onsets. Double accent was more common in trisyllabic words whose initial syllable has aspirated onsets and final accent was assigned more often in bisyllabic words with aspiration on final onsets. These results suggest that phonetics-based constraints as well as phonologically-grounded constaints are active in NKK accentuation of new words. Therefore, although penultimate accent is dominant in NKK accentuation due to constraints which identify the penult as a default position, non-penultimate accent is also present as a variant by speakers who are more sensitive to the association between aspiration and higher F0 in accentuation of new words.

Section 7.2 presents linguistic background about phonetic effects of aspiration on adjacent vowels and lexical accent patterns in NKK. Section 7.3 presents an experimental study to examine phonetic effects on NKK accentuation in novel words. It tested whether aspiration on onsets would influence accentuation of novel words. Section 7.4 discusses phonetic characteristics of accent patterns induced by aspiration. Section 7.5 discusses the results and Section 7.6 concludes.

7.2 Aspiration Effect

The relationship between aspiration on obstruent consonants with fundamental frequency (F0) of the following vowel in Korean has been discussed in many studies (Han & Weitzman 1965; Silva 1992, 2006; S-A Jun 1993; M-R Kim 2000; M-R Kim *et al*. 2002; M-R. Kim and Duanmu 2004; Kenstowicz and Park 2006; and many other studies). F0 was significantly higher after aspirated consonants than after lax consonants in Jeonnam Korean and Seoul Korean, which are not pitch accent languages (S-A Jun 1993; M-R. Kim and Duanmu 2004; Silva 2006). Voice onset time (VOT) has been assumed to be a robust phonetic cue for the laryngeal contrast among lax, tense and aspirated stops in Korean in that VOT is shortest for Korean tense stops (6-18ms), longer for lax (20-60ms), and longest for aspirated (100-115ms), whereas F0 has been assumed to be a redundant feature for the phonation type contrast (Lisker & Abramson 1964; C-W. Kim

1965; Silva 1992) On the other hand, M-R Kim (2000) proposed that the consonantal effect on F0 is tonal, arguing that aspirated/tense stops, affricates and fricatives formed a H tonal pattern in terms of the F0 contour following constriction release, whereas lax stop/affricates and sonorants formed a LH tonal pattern.

The aspiration effect on pitch (F0) of the following vowel also holds in Kyungsang Korean (e.g. Kenstowicz and Park 2006). For the sake of the comparison of dialectal differences, Lee and Jongman (2010) studied production of the three phonation types in Seoul Korean and in Kyungsang Korean (KK). They found that F0 was highest after aspirated stops, intermediate after tense stops, and lowest after lax stops both in Seoul Korean and in KK. Contra the previous assumption, VOT differences were not informative enough to distinguish the phonation types in Seoul Korean because of overlap between lax and aspirated stops, whereas VOT values were distinct enough for the contrast in KK. On the other hand, F0 itself could not serve as a robust cue for the contrast in KK due to overlap in F0 values between tense in the low tone condition and lax in the high tone condition, and between tense in the high tone and aspirated in the low tone condition. Although the consonantal effect on F0 was not a good source for the three way phonation type contrast in KK, the F0 difference was informative for the two way contrast between aspirated and lax stops in KK speakers' production just as in Jeonnam Korean and Seoul Korean.

However, aspirated onset consonants are not necessarily associated with a high tone pitch accent in NKK lexical accentuation, as illustrated in the examples in (60).

(60) NKK lexical accentuation irrelevant to onset quality

i) *Double accent*

t^húgú 'knight's helmet'

c^hébí 'preparation'

hə́p^há 'lung'

k^hɨdári 'long leg'

ii) *Final accent*

c^himá 'skirt'

p^hodó 'grape'

t^hacó 'ostrich'

koc^hú 'pepper'

iii) *Penult accent*

c^hə́ma 'eaves'

p^hógi 'head of vegetables'

púp^hi 'bulk'

p^hodégi 'baby blanket'

Bisyllabic/trisyllabic words beginning with aspirated consonants can have all the different possible patterns (final accent as well as double accent and penultimate accent) (e.g. p^hodó, t^húgú, p^hógi; k^hɨdári, p^hodégi). In addition, aspirated onset consonants in the final syllable do not necessarily attract accent (koc^{h}ú vs. púp^{h}i). The aspiration effect on pitch (F0) of the following vowel does not play a role in NKK lexical accent patterns.

Without underlying information about the accent position, it is plausible that NKK speakers utilize the phonetic effect of prevocalic consonants when accenting newly adopted words. Furthermore, it is also possible that NKK speakers would perceive the tonal difference between aspirated and lax stops as a cue for lexical accent when they encounter new words produced in other dialects (e.g. Seoul Korean, Jeonnam Korean). Given all these possibilities, onset-sensitive accentuation is hypothesized for accent patterns of newly adopted words in NKK. In order to test the consonantal effects on pitch accent, an experimental study was performed using novel words.

7.3 Experiment 3

A hypothesis which guided this study is that accentuation will be sensitive to onsets (c.f. Gordon 2005). Based on previous literature showing that F0 was realized significantly higher after aspirated consonants than after lax consonants in Korean, the hypothesis predicts that aspirated onset consonants should be more likely to attract accent than other consonants in novel words, due to the association between aspiration and F0.

7.3.1 Participants and Procedure

The participants were the same as those in the first experiment presented in Chapter 3: 30 NKK speakers, ranging in age from 29 to 54 years (mean 39), 15 male and 15 female. The current experiment was done right after Experiment 2 with the same procedure.

7.3.2 Materials

The test word set consisted of novel words with four possible combinations of CV syllables for bisyllabic words and for trisyllabic words: NN, AN, NA, AA, NNN, ANN, AAN, NAN (N:CV with a non-aspirated onset; A:CV with an aspirated onset) (e.g. *tapa, t^{h}apa, tap^{h}a ,t^{h}ap^{h}a, kitici, k^{h}itici, k^{h}it^{h}ici, kit^{h}ici*). Only obstruent consonants were used for onsets such as *p, t, k* (e.g. *pakapa*, *citiki*) in order to test aspiration effects, because only obstruents have the laryngeal contrast. A single vowel, either *a* or *i*, was used within a word to avoid the possibility that vowel quality might influence accent placement: for example, only the vowel *a* was used as in a word *tapa* or the vowel *i* was used as in a word *kitici*. Four items were included for each combination. In total, 32 words were created: 16 bisyllabic words and 16 trisyllabic words.

(61) Novel stimuli (N=nonaspirated onset; A=aspirated onset)

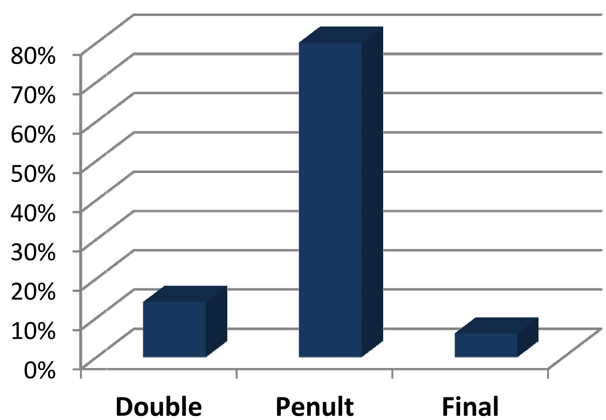
	Type 1	*Type2*	*Type 3*	*Type 4*
Bisyllabic:	NN	AN	NA	AA
	tapa	t^{h}apa	tap^{h}a	t^{h}ap^{h}a
	kapa	k^{h}apa	kap^{h}a	k^{h}ap^{h}a
	piki	p^{h}iki	pik^{h}i	p^{h}ik^{h}i
	cipi	c^{h}ipi	cip^{h}i	c^{h}ip^{h}i

Trisyllabic:	NNN	ANN	AAN	NAN
	pakapa	p^{h}akapa	p^{h}ak^{h}apa	pak^{h}apa
	takapa	t^{h}akapa	t^{h}ak^{h}apa	tak^{h}apa
	kitici	k^{h}itici	k^{h}it^{h}ici	kit^{h}ici
	citiki	c^{h}itiki	c^{h}it^{h}iki	cit^{h}iki

7.3.3 Results

A total of 951 tokens were collected from 32 test words with 30 subjects (excluding 9 tokens with production errors). 480 tokens were bisyllabic and 471 tokens were trisyllabic. Chapter 3 showed that penultimate accent was predominant in words with all light syllables: 82% of bisyllabic words and 81% of trisyllabic light-syllabled words were given penultimate accent. Nevertheless, the hypothesis predicts that non-penultimate accent would be more frequent in the combinations having aspirated onsets in non-penultimate syllables such as NA, ANN than in other combinations such as NN, AN, NNN, NAN, because aspirated onset consonants would attract accent. The overall results of bisyllabic words are presented in Figure 45. Penultimate accent was found most frequently (80%), double accent was assigned considerably less (14%), and final accent was rarely found (6%) in bisyllabic words.

Figure 45 Overall results in bisyllabic words



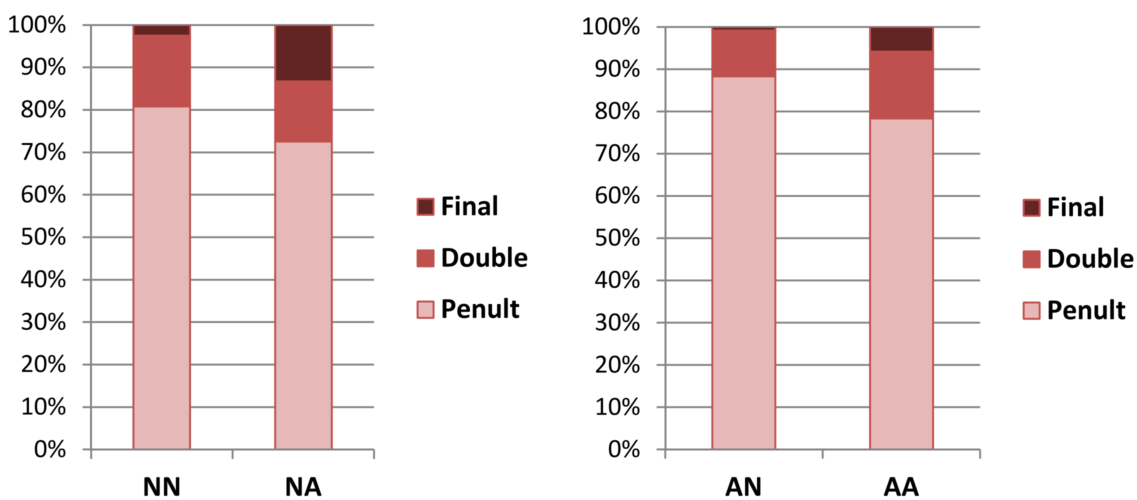
Accent distribution of each combination also showed the patterns similar to that of the overall results. As shown in Table 18, penultimate accent was the most common across all the combinations: 81% in NN; 88% in AN; 72% in NA; 78% in AA. Double accent was given around 15% of the time in most combinations: 17% in NN; 11% in AN; 14% in NA; 16% in AA. Final accent is very rare since few words were given final accent in most combinations except in NA combinations.

Table 18 Results according to syllable combinations in bisyllabic words

	Double	Penult	Final	Total
NN	20 (17%)	**97 (81%)**	3 (2%)	120 (100%)
AN	13 (11%)	**106 (88%)**	1 (1%)	120 (100%)
NA	17 (14%)	**87 (72%)**	16 (13%)	120 (100%)
AA	19 (16%)	**94 (78%)**	7 (6%)	120 (100%)
Total	69 (14%)	**384 (80%)**	27 (6%)	480 (100%)

To compare the distribution of accent among the word types (NN; AN; NA; AA), statistical analyses using Pearson's Chi-square were performed. The results showed that the distribution of accent in AN combinations was not different from that of NN ($\chi^2(2)=2.88$, p=.24), which suggests that aspirated consonants in the onset of the initial syllable do not disturb the typical accentuation of NN. However, the distribution of accent was significantly different between NN and NA ($\chi^2(2)=9.68$, p=.008) and between AN and AA ($\chi^2(2)=6.35$, p=.04). Final accent is more frequent in NA than in NN and penultimate accent is less frequent in AA than in AN, due to the increase of double accent and final accent. This suggests that onset aspiration of final syllables is a factor to induce non-penultimate accent in words with light syllables. Figure 46 illustrates the different accent distributions between NN and NA, and between AN and AA combinations.

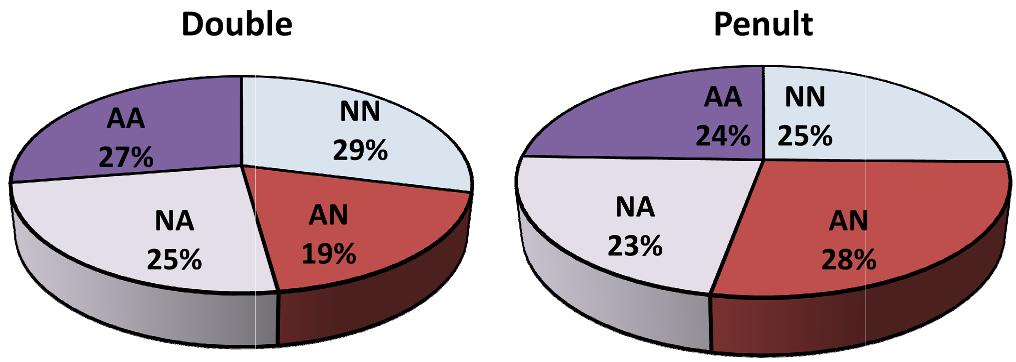
Figure 46 Accent distribution in bisyllabic words



As shown in the graphs in Figure 46, final accent was assigned more frequently in NA and AA combinations.

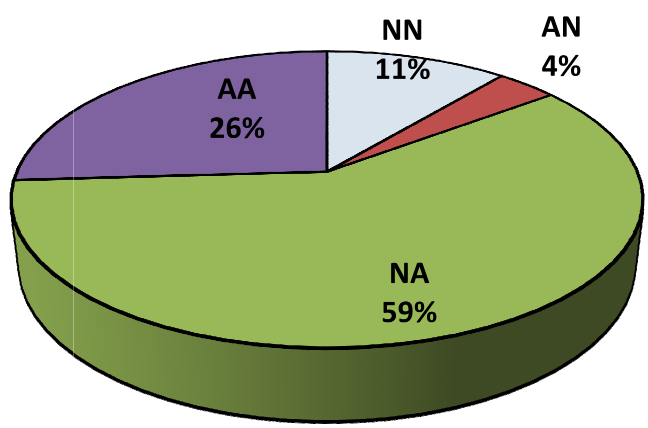
The distribution of each accent pattern is given in the graphs in Figure 47 and Figure 48. Double accent was distributed evenly for all the combinations: 29% (NN); 19% (AN); 25%(NA); 27%(AA). This suggests that double accent placement is not influenced by the quality of onset consonants in bisyllabic words. Penultimate accent also showed similar results with double accent assigned relatively equally for all the combinations: 25% in NN; 28% in AN; 23% in NA; 24% in AA. This suggests that onset quality did not matter in the assignment of double accent and penultimate accent in bisyllabic words.

Figure 47 Even distribution of double accent and of penultimate accent in bisyllabic words



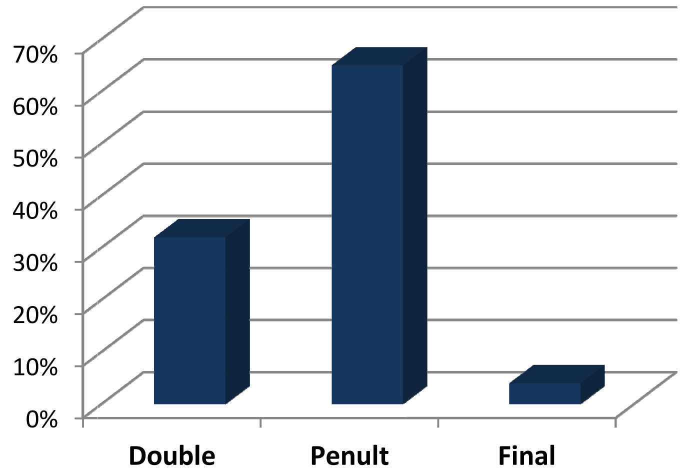
On the other hand, as shown in Figure 48, the distribution of final accent was skewed to NA combinations: 59% of final accented words were NA combinations, 26% were AA combinations and the rest 15% were NN and AN combinations. Final accent was fairly rare in words having only light syllables: only 27 words out of 480 words (6%) were given final accent. However, out of 27 final accented words, 23 words (85%) had aspirated onsets in the final syllable (NA; AA). The distribution of final accent clearly demonstrates that final accent was not assigned randomly but rather was affected by the phonetic quality of onsets in final syllables, despite avoidance of final accent.

Figure 48 Uneven distribution of final accent in bisyllabic words



The overall results for trisyllabic words were also similar to the patterns of bisyllabic words, as shown in Figure 49. Penultimate accent was most frequent (65%), double accent was less common (32%), and final accent was least frequent (4%).

Figure 49 Overall results in trisyllabic words



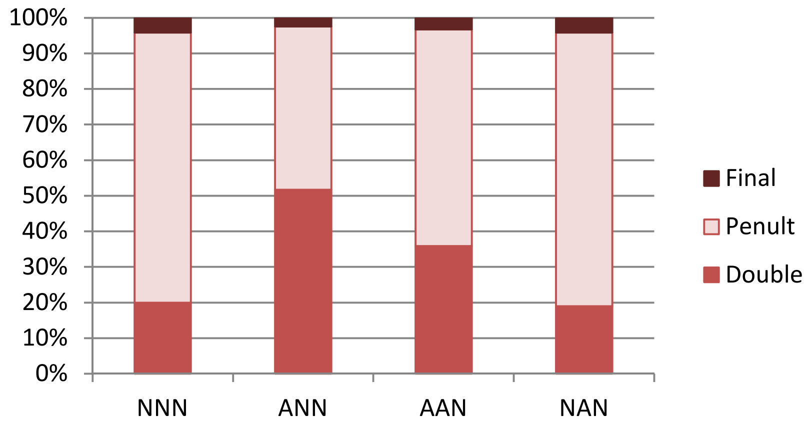
However, Table 19 shows that the patterns for particular syllable combinations were different from those of bisyllabic words. When a word-initial consonant was aspirated, double accent was more likely: 52% of ANN combinations were given double accent, which contrasted with the dominant penultimate accentuation in other combinations. The patterns in the other combinations complied with overall patterns in that penultimate accent was most common.

Table 19 Results according to syllable combinations in trisyllabic words

	Double	Penult	Final	Total
NNN	24 (20%)	**90 (76%)**	5 (4%)	119 (100%)
ANN	**62 (52%)**	55 (46%)	3 (3%)	120 (100%)
AAN	41 (36%)	**70 (61%)**	4 (3%)	115 (100%)
NAN	22 (19%)	**90 (77%)**	5 (4%)	117 (100%)
Total	149 (32%)	**305 (65%)**	17 (4%)	471 (100%)

Figure 50 clearly illustrates that accent distribution was not consistent across syllable combinations in trisyllabic words.

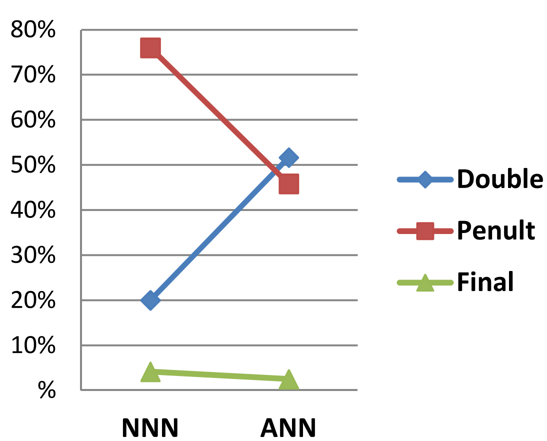
Figure 50 Distribution of accent according to syllable combinations in trisyllabic words

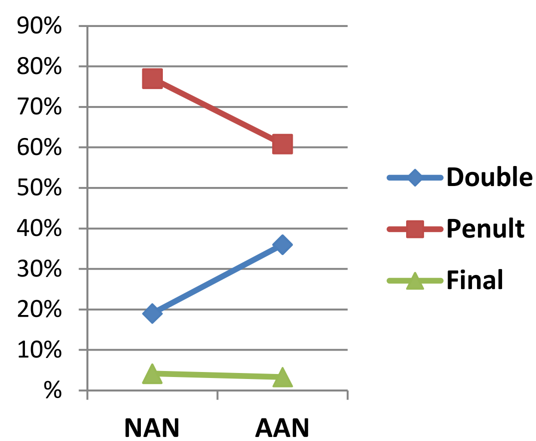


Double accent was more frequent in ANN and AAN combinations than in other combinations. Penultimate accent was predominant in NNN and NAN combinations but not in ANN and AAN combinations, whereas final accent was rare across all the combinations.

The graphs in Figure 51 show that the onset quality of word-initial syllables influenced accent placement, seen by comparing accent patterns between NNN and ANN, and between NAN and AAN.

Figure 51 Accent placement in NNN vs. in ANN; in NAN vs. in AAN





Double accent was more common in ANN and AAN than in NNN and NAN, while penultimate accent was less common in ANN and AAN than in NNN and NAN. Pearson's Chi-square analyses confirmed that the differences between NNN and ANN and between NAN and AAN were statistically significant ($\chi^2(3)=26.74$, $p<.001$; $\chi^2(3)=8$, $p<.05$).

To summarize, penultimate accent was preferred both in bisyllabic words and in trisyllabic words. However, the predominant penultimate accentuation was disturbed in trisyllabic words when the word began with aspirated onset consonants: only 46% of trisyllabic words with word-initial aspirated onsets received penultimate accent while 76% of words without aspiration onsets received penultimate accent. As for bisyllabic words, aspirated onsets in word-initial syllables did not bring about a change in accent patterns. This is not surprising because word-initial aspirated onsets in bisyllabic words do not create a conflict with the preferred penultimate accent position. However, 13% of words containing aspirated onsets in the final syllable received final accent. The difference between final accent patterns in NA and in NN was statistically significant since final accent was very rare in words with all light syllables (2%).

These results suggest that NKK speakers are sensitive to onset quality when they accent novel words. Although the consonantal effect on accentuation was not categorical, the phonetic effect was important enough to serve as a factor in deriving variation in accent patterns in newly adopted words. For example, variations of accent patterns found in exisitng loanwords, as in *nepk^{h}ín* 'napkin' with final accent and *k^{h}épt^{h}ín* 'captain' with double accent, might be ascribed to this phonetic effect. Some speakers accent new words more by following a favored position, which will produce dominant penultimate accentuation, while other speakers are more sensitive to the association between aspiration and higher F0 and apply it to accentuation of new words.[40]

7.4 Phonetic Characteristics of Onset-Sensitive Accent

Acoustic analyses were conducted to examine phonetic characteristics of accented vowels and unaccented vowels of ANN and of NNN. This study was initiated to confirm accuracy of the accent judgments because the results of the previous section were based on aural judgments. Furthermore, this study also investigated what other phonetic cues besides F0 (e.g. F0 contour) were available for the accentual contrast between penultimate accent and double accent in cases when F0 was higher in low toned vowels (V1 following aspirated consonants in ANN with penultimate accent) than high toned vowels (V1 following unaspirated consonants in NNN with double accent).

This study was guided by the hypothesis that phonetic cues are sufficient for the contrast between accented vowels and unaccented vowels of ANN and of NNN if misperception was not involved in the accent judgments. Because F0 itself cannot serve as a cue for the accentual contrast in some cases (e.g. unaccented V1 following aspirated consonants vs. accented V1 following unaspirated consonants), F0 contours generated from all the three vowels in a word could be an additional informative cue. We predict that ANN with double accent will manifest different phonetic patterns (e.g. F0 values or F0 contour) from ANN with penultimate accent just as NNN with double accent does from NNN with penultimate accent. Average F0 and F0 contours of the vowel in each syllable in ANN/NNN with double accent (**High** tone category since the target initial syllable has a high tone) and in ANN/NNN with penultimate accent (**Low** tone category since the target initial syllable has a low tone) were compared using a repeated-measures analysis of variance (ANOVA). The dependent variable was the mean F0 of the vowel in each syllable of the target word. The independent variables were onset type of the first syllable (aspirated vs. unaspirated) and tone of the first syllable (high vs. low).

[40] 7 out of 30 NKK subjects (23%) belong to a rigid onset-sensitive group: they followed onset-sensitive accent patterns (e.g. double accent for ANN/AAN forms) in 75% of the time (42/56), whereas 8 out of the NKK subjects (27%) belong to a penultimate accent group: they showed robust preferences for penultimate accent regardless of onset quality (92%; 118/128). The remaining group of the NKK subjects (15/30) showed penultimate accent in 56% of the time (134/239) and onset sensitivity in 51% of the time (61/120).

7.4.1 Materials

A total of 464 tokens were analyzed: 4 ANN words and 4 NNN words with 2 repetitions recorded by 15 male and 15 female speakers.[41]

(62) Target item list

ANN	NNN
c^hitiki	citiki
k^hitici	kitici
p^hakaba	pakapa
t^hakaba	takapa

7.4.2 Measurement

The following measurement was taken: average F0 of the vowels of the target word's three syllables at four points (1/4, 2/4(midpoint), 3/4, 4/4(offset)). F0 was measured by selecting the span of the vowel in the acoustic display and allowing Praat to automatically calculate the desired value. Then the following values were compared: F0 contours and average F0 in V1 (High) of ANN words with double accent and in V1 (Low) of ANN words with penultimate accent. In addition, F0 contours and average F0 in V1 (High) of NNN words with double accent and in V1 (Low) of ANN words with penultimate accent were also compared.

7.4.3 Results

The patterns of ANN and NNN were compared according to tone category (High vs. Low) and onset type of the first syllable (aspirated vs. unaspirated) using a repeated-measures analysis of variance (ANOVA). Average F0 values of V1 in NNN and in ANN measured at the midpoint are presented in Table 20. In addition, the comparison of mean F0 between low toned V1 (V1 of tokens with penultimate accent (LHL)) and high toned V1 (V1 of tokens with double accent (HHL) is also shown in Table 20.

Table 20 Compared mean F0 of V1 midpoint in NNN vs. in ANN; in LHL(penult) vs. in HHL(double)

	Male		Female		N
	Mean F0 (Hz)	S.D	Mean F0 (Hz)	S.D	
NNN	127.7	21.2	217.1	21.2	30
ANN	162.2**	24.1	275.3**	37.8	30
LHL	128.9	25.4	228.4	50.1	26
HHL	164.4*	24.5	280.2**	40.2	26

* indicates that the differences are significant at α=.05
** indicates that the differences are significant at α=.001

[41] One male speaker produced only one repetition while the other 29 participants repeated all tokens twice and 8 items were excluded from the analyses due to production errors or devoiced vowels.

As shown in Table 20, the aspiration effect on F0 values was significant: the F0 of V1 in NNN is consistently lower than the F0 of V1 in ANN, for both male and female speakers (male: $F(1,14)=88.4$, $p<.001$; female: $F(1,14)=41.8$, $p<.001$). The tonal effect on F0 values was also significant: the F0 of V1 in LHL (ANN/Low; NNN/Low) is significantly lower than in HHL (ANN/High; NNN/High) (male: $F(1,11)=21.7$, $p=.001$; female: $F(1,13)=36.5$, $p<.001$).

Table 21 presents average F0 values of V1 measured at the midpoint according to each tone (H;L) for each word type (NNN;ANN).

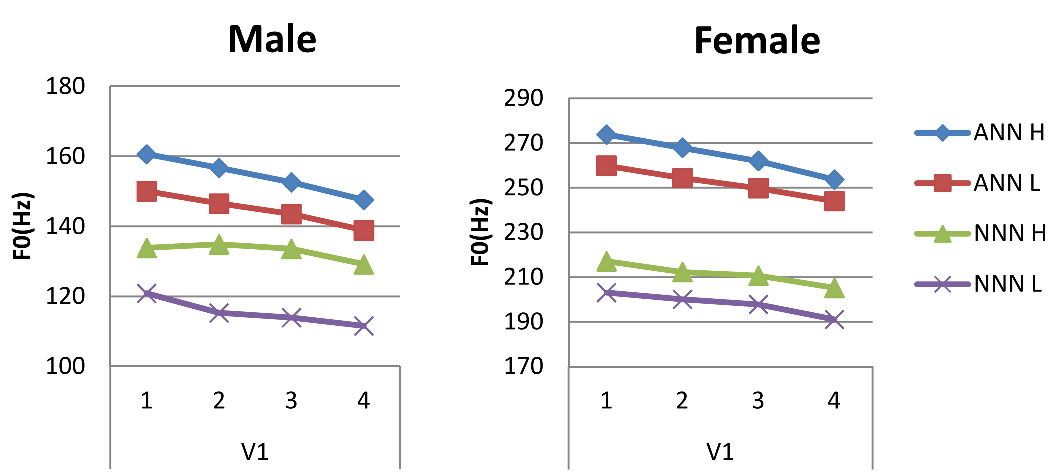
Table 21 Comparison of mean midpoint F0 between low toned V1 and high toned V1 for each word type

	Male				Female			
	NNN		ANN		NNN		ANN	
	L	H	L	H	L	H	L	H
Mean F0 (Hz)	114.5	134.9	143.5	159.6*	196.1	212.3	254.4	267.9**
S.D	14.4	22.9	22.0	29.1	15.0	7.7	35.2	36.5
N	15	3	13	13	15	8	12	14

* indicates that the differences are significant at α=.05
** indicates that the differences are significant at α=.001

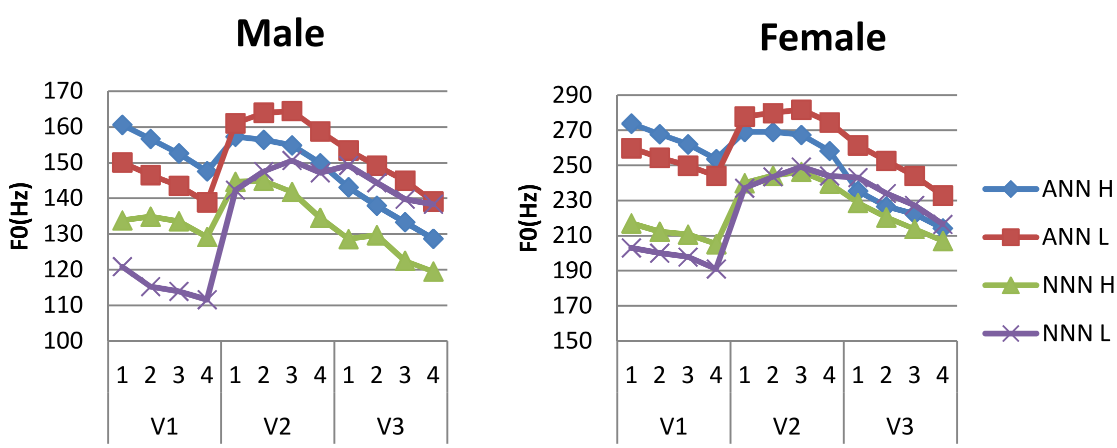
The tonal effect on the F0 of V1 in ANN was statistically significant (Male: $F(1,9)=5.98$, $p=.04$; Female: $F(1,10)=61.4$, $p<.001$). As shown in Table 21, F0 values were consistently lower in ANN (Low) than in ANN (High), even if F0 begins with relatively higher F0 in ANN (Low). The tonal effect was also found in NNN groups: F0 values were consistently higher in NNN (High) than in NNN (Low). In terms of F0 values, four distinct levels were generated for each category: NNN (Low) < NNN (High) < ANN (Low) < ANN (High). The following figures illustrate distinct F0 levels of V1 for each category.

Figure 52 Distinct F0 values of V1 for ANN (High; Low) and NNN (High; Low)



In terms of distinction between low toned V1 in ANN and high toned V1 in NNN, F0 of the vowel might not be sufficient for the tonal contrast since the F0 of low toned V1 in ANN was consistently higher than that of high toned V1 in NNN. Thus, I compared F0 contours of the four categories. The following figures in Figure 53 display F0 contours based on average F0 measured at four points of three syllables in target words. Male speakers and female speakers showed similar patterns for pitch contour of each category.

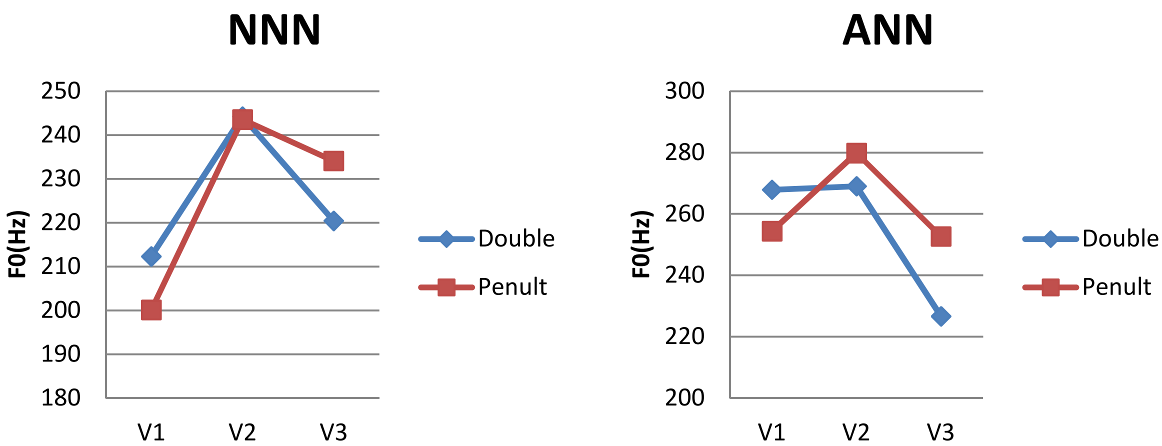
Figure 53 Pitch contour patterns of NNN (H:double; L:penult) and ANN (H:double; L:penult)



For both male and female speakers, ANN with double accent(ANN_H) began with the highest pitch among the four categories. Yet the pitch peak of the second syllable was not higher than that of the first syllable. The pitch fall in the third syllable was deep as F0 was lower than that of ANN(Low) and NNN(Low). ANN with penultimate accent (ANN_L) began at a lower pitch than ANN (High) but the pitch peak was higher than that of ANN (High) in the second syllable. The pitch fall in the third syllable was less steep, so the pitch was maintained higher than that of ANN (High). The pattern of ANN (Low) was also distinct from NNN(High) in that the pitch fall was relatively gentle and the pitch level was fairly high in the third syllable. On the other hand, the contour shapes of NNN (High) were similar to those of ANN (High) except the relative low F0 in the first syllable.

The pitch contour shapes in Figure 54 generated from average F0 at the midpoint of each syllable more clearly demonstrate the distinct patterns of each category. Only the graphs for female speakers are presented, since the patterns of male speakers were consistent with them.

Figure 54 Pitch contour patterns of NNN and ANN: Double accent vs. Penult Accent



The pitch peak was on the second syllable in both NNN(High) and NNN(Low), but the high tone began higher than the low tone in the first syllable. Recall that NNN (High) words contained only lax onsets with double accent (HHL) while NNN (Low) words had penultimate accent (LHL). While the final syllable in both categories had a low tone, the pitch fall was steeper in NNN (High) than in NNN (Low). The patterns shown in NNN novel words were consistent with D-M Lee (2008a)'s description of the phonetic characteristics of NKK tone patterns based on trisyllabic existing words beginning with sonorant onsets: the pitch of the second syllable was higher than that of the first syllable in double accent (high category), and no deep fall was found after the peak in the third syllable in penultimate accent (low category).[42] On the other hand, when a word began with an aspirated onset, the pitch was elevated in the beginning, forming a pitch plateau in ANN(High), which resulted in a different contour shape from NNN (High). The pitch fell steeply in ANN (High), while the pitch fall was much less steep in ANN (Low). Therefore, the pitch contours of each category can be characterized as follows:

(63) Characteristics of pitch contours
ANN (High): pitch plateau with steep pitch fall on the third syllable
ANN (Low): pitch peak on the second syllable and less steep pitch fall on the third syllable
NNN (High) : pitch peak on the second syllable but steep pitch fall on the third
NNN (Low) : steep rise to the pitch peak on the second syllable and gentle fall on the third

Both ANN (Low) and NNN (High) have a pitch peak on the second syllable, but ANN (Low) has a gentle pitch fall on the third syllable, characteristic of penultimate accent, whereas NNN (High) has a steep pitch fall of double accent on the third syllable. Even though the F0 of ANN (Low)

[42] D-M Lee did not use words beginning with an aspirated consonant in his study, so the patterns of ANN existing words were not found yet.

began higher than that of NNN (High), the distinct pitch contour shape of penultimate accent indicated that the initial syllable of ANN (Low) had a low tone.

To sum up, the phonetic qualities of ANN/NNN (High) were distinct from those of ANN/NNN (Low): i) average F0 values of the vowel in the initial syllable in ANN/NNN (High) were significantly higher than those in ANN/NNN (Low); ii) the pitch contour of ANN (High) was distinct from that of ANN (Low): ANN (High) began with a pitch plateau at a higher F0 and ended with steep pitch fall, whereas ANN (Low) began with a lower F0 but had a steep pitch rise to the second syllable; iii) the pitch contour of NNN (High) was also different from that of NNN (Low): the pitch peak was on the second syllable in both NNN(High) and NNN(Low), but the high tone began higher than the low tone in the first syllable and the pitch fall was steeper in NNN (High) than in NNN (Low). These results suggest that the phonetic information on the surface was sufficient to cue the distinction between ANN (High) and ANN (Low). Therefore, we can assume that accent was accurately identified, and that accentuation was indeed affected by onset aspiration.

7.5 Discussion

This study provided evidence that aspiration plays a role in NKK accentuation of novel words, even though aspirated onsets are not necessarily associated with a high tone pitch accent in NKK lexical accentuation. Acoustic analyses showed that phonetic patterns of vowels with aspirated onsets, judged as accented, were clearly distinct from those of the vowels judged as unaccented, confirming that phonetic information was sufficient for the accentual contrast. However, the onset-sensitive accentuation was not categorical, instead being associated with variation in the accent patterns of newly adopted words. Therefore, a phonetics-based source also serves as one of the factors to motivate NKK accent placement of novel words; for example, NKK speakers might perceive higher F0 as signaling accent because F0 is utilized for the tonal contrast in NKK.

7.5.1 Onset-sensitive patterns vs. onset-insensitive patterns

Penultimate accent was preferred both in bisyllabic words and in trisyllabic words when a word consisted of only light syllables. However, predominant penultimate accentuation was disturbed in trisyllabic words when a word beginning with aspirated consonants was included: 46% of trisyllabic words with word-initial aspirated onsets were given penultimate accent, while 76% of words without aspirated onsets were given penultimate accent. In bisyllabic words, aspiration of the initial onset did not influence accent placement of novel words. I assume that this is because accent attraction of word-initial aspirated onsets does not create conflict with the preferred penultimate accent in bisyllabic words. However, when aspirated onsets were in the final syllable, final accent increased to 13%, which was statistically significant since only 2% of the words with all light syllables had final accent. The results suggest that onset quality was relevant to accentuation of novel words. Some speakers accent new words more restictively following favored position, which will produce dominant penultimate accentuation, while some speakers would be more sensitive to phonetic quality and produce onset-sensitive accentuation in new words.

7.5.1.1 Dominant penultimate accentuation

The predominant penultimate accentuation in trisyllabic words is explained as default accentuation, proposed by N-J Kim (1997). The following constraints are relevant to the penultimate accent assignment, repeated from (32) in 6.3.1:

(64) Constraints for default penultimate accent (based on N-J Kim 1997)
*..H..]$_{PW}$: A prosodic word must have a high tone.
ALIGN-R: Align the right edge of a high tone with the right edge of a word.
NONFINALITY : No high tone on the final syllable of a word is allowed.
*H: No high tone is allowed

The penultimate accent placement is explained by the ranking of the constraints presented in (65).

(65) *..H..]$_{PW}$ >> NONFINALITY >> ALIGN-R>>*H

(66) Default penultimate accentuation

Input: /takapa/	*..H..]$_{PW}$	NONFINALITY	ALIGN-R	*H
a. takapa	*!			
b. takapá		*!		*
c. tákapa			**!	*
d. tákápa			*	*!*
☞e. takápa			*	*

Novel words are assumed to have no underlying high tone accent. A high tone must be assigned to satisfy the constraint *..H..]$_{PW}$. The constraint NONFINALITY ranked higher than ALIGN-R forces the penultimate syllable (66e) to have a high tone accent. Antepenultimate accent (66c) loses because it violates ALIGN-R twice. Double accent (66d) cannot make an optimal output because it violates *H twice.

7.5.1.2 Onset-sensitive accentuation

When a word-initial consonant was aspirated, double accent was more likely: 52% of ANN combinations were given double accent, which contrasted with the dominant penultimate accentuation in other combinations. The constraints in (64) identifying the default accent placement are not satisfactory to predict this context-sensitive patterning. A constraint which prohibits syllables with aspirated onset consonants from having low tone is motivated to account for the accent placement. The association of aspiration with a high tone is also attested cross-linguistically. For example, S-H Lee (2008) showed that a high tone is only allowed for syllables with aspirated onsets in Mulao, a language spoken in the Guangxi Zhuang Autonomous Region in southwest China (e.g. [p^{h}ó^{44}] 'bed' H tone vs. *[p^{h}o] L tone). Hsieh & Kenstowicz (2008) also found that in Lhasa Tibetan, words borrowed from both English and Mandarin are assigned tone according to the laryngeal features of the onset: H tone when the onset is an obstruent and L tone when the onset is a sonorant. Burmese is another case where coda consonants restrict tonal realization: for instance, syllables with a glottal stop coda are not allowed to have a high tone ([k^{h}aʔ] ‘to draw off’ L tone vs. *[k^{h}áʔ] H tone), S-H Lee argues that “consonant-tone interaction results from the requirement that all segments in the output must be associated to a tone, which is ensured by the markedness constraint RootNode →T”(p.129). Therefore, all segments are under a dependency relationship with tone. He proposes a markedness constraint *[+SpreadGlottis]/Low (67) for the association of aspiration with a high tone, which penalizes aspirated consonants that form a dependency relationship with a low tone in the output.

(67) *[+SpreadGlottis]/Low (S-H Lee 2008, p141)
No consonant specified as [+SpreadGlottis] associates to L tone in the output.

With respect to the preference of double accent for ANN in NKK, the constraint *[+SpreadGlottis]/Low triggers a high tone for syllables with aspirated onset consonant. However, a question arises: why is double accent assigned rather than antepenultimate accent when only the antepenultimate syllable contains an aspirated onset consonant? I propose that the preference for double accent rather than antepenultimate accent in ANN results from conspiracy of the constraints Align-R and *[+SpreadGlottis]/Low. The following tableau illustrates the case when an aspirated onset consonant attracts a high tone pitch accent.

(68) Onset-sensitive accentuation

Input: /t^{h}akapa/	Nonfinality	Align-R	*[+SpreadGlottis]/Low	*H
a. t^{h}akápa		*	*!	*
b. t^{h}akapá	*!		*	*
c. t^{h}ákapa		**!		*
☞d. t^{h}ákápa		*		**

The candidates (68a) and (68b) with penultimate accent/final accent lose because of a violation of *[+SpreadGlottis]/Low and of Nonfinality respectively. The candidate (68c) with antepenultimate accent also loses because it violates Align-R twice, although it satisfies *[+SpreadGlottis]/Low. Because double accent is represented with just a single H but shared by two syllables (Chapter 2, p. 9), the candidate (68d) violates Align-R once, while it violates *H twice. Therefore, the double accented candidate (68d) makes an optimal output without critical violations. If a NKK speaker produced accent sensitively to the association of aspirated consonants with a high tone, the constraint *[+SpreadGlottis]/Low must outrank the constraint *H in his/her grammar. If a NKK speaker preferred penultimate accent regardless of onset quality, his/her grammar must demote the constraint *[+SpreadGlottis]/Low lower than the constraint *H.

(69) Onset-insensitive accentuation

Input: /t^{h}akapa/	Nonfinality	Align-R	*H	*[+SpreadGlottis]/Low
☞a. t^{h}akápa		*	*	*
b. t^{h}ákápa		*	*!*	

As shown in (69), penultimate accent (69a) is a winner despite the word-initial aspirated consonant because the constraint *[+SPREADGLOTTIS]/LOW is not crucial in the accentuation in this grammar.

7.5.2 GLA Simulation

Chapter 6 showed that a simulation assuming the Gradual Learning Algorithm (GLA)(Boersma1997; Boersma & Hayes 2001) predicted double accent with a lower frequency than the actual number of double accents found in novel words. I assumed that this mismatch was because the simulation ignored the phonetic effects. This chapter has shown that phonetic qualities (e.g. aspiration; F0) indeed serve as another factor in accent variation in novel words, though the phonetic effects on accentuation were not categorical: syllables with aspirated onset consonants were more likely to carry a high pitch accent.[43] I tested whether the GLA successfully learn the phonetic effects on NKK accent patterns in novel words. To capture the aspiration effects, the GLA was run including the constraint *[+ SPREADGLOTTIS]/LOW, repeated in (70), to the input file, in addition to other relevant accent markedness constraints discussed so far.

(70) *[+SPREADGLOTTIS]/LOW
No consonant specified as [+SPREADGLOTTIS] associates to L tone in the output.

The training data set consists of 80 pairs of given underlying forms together with surface forms, as listed in (71): four different underlying forms of trisyllabic words according to syllable combinations: CV.CV.CV, C^{h}V.CV.CV, CV.C^{h}V.CV, and C^{h}V.C^{h}V.CV; for each underlying form, five different accent patterns were assigned: no accent, initial accent, penultimate accent, final accent, and double accent; and for each underlying form, four different items were included; e.g., for a CV.CV.CV form, *takapa*, *pakapa, citiki,*and *kitici* were included. The training data were as follows:

(71) Training data

underlying forms	|cv.cv.cv|	|c^{h}v.cv.cv|	|cv.c^{h}v.cv|	|c^{h}v.c^{h}v.cv|
	takapa	t^{h}akapa	tak^{h}apa	t^{h}ak^{h}apa
	pakapa	p^{h}akapa	pak^{h}apa	p^{h}ak^{h}apa
	kitici	k^{h}itici	kit^{h}ici	k^{h}it^{h}ici
	citiki	c^{h}itiki	cit^{h}iki	c^{h}it^{h}iki

[43] In addition, Chapter 3 demonstrated that vowel quality was also a factor of accent variations in novel words: double accent was more likely in words beginning with a high vowel. In order to account for the vowel effects on accent patterns, a constraint which prohibits a low tone from the tone bearing unit with high pitch on the surface form should be necessary. To determine whether such a constraint is universal or language-specific, more research is required. I would rather not attempt an analysis for the vowel effects and leave them for future studies.

surface forms	/cv́.cv.cv/	/c^{h}v́.cv.cv/	/cv́.c^{h}v.cv/	/c^{h}v́.c^{h}v.cv/
	/cv.cv́.cv/	/c^{h}v.cv́.cv/	/cv.c^{h}v́.cv/	/c^{h}v.c^{h}v́.cv/
	/cv.cv.cv́/	/c^{h}v.cv.cv́/	/cv.c^{h}v.cv́/	/c^{h}v.c^{h}v.cv́/
	/cv́.cv́.cv/	/c^{h}v́.cv́.cv/	/cv́.c^{h}v́.cv/	/c^{h}v́.c^{h}v́.cv/
	/cv.cv.cv/	/c^{h}v.cv.cv/	/cv.c^{h}v.cv/	/c^{h}v.c^{h}v.cv/

The input frequencies were based on the results in Experiment 3, reported on in 7.3, as repeated in (72).

(72) Input frequencies:

Word type		Penult	Final	Double	Initial
CVCVCV	takapa	21	2	7	0
	pakapa	26	1	2	0
	kitici	21	1	8	0
	citiki	22	1	7	0
C^{h}VCVCV	t^{h}akapa	16	1	13	0
	p^{h}akapa	18	1	11	0
	k^{h}itici	7		23	0
	c^{h}itiki	14	1	15	0
CVC^{h}VCV	tak^{h}apa	27	1	2	0
	pak^{h}apa	20	1	9	0
	kit^{h}ici	20	2	8	0
	cit^{h}iki	24	1	5	0
C^{h}VC^{h}VCV	t^{h}ak^{h}apa	21	1	7	1
	p^{h}ak^{h}apa	20	1	9	0
	k^{h}it^{h}ici	17	1	12	0
	c^{h}it^{h}iki	20	1	9	0

In a representative run, the algorithm generated the following ranking probabilities in (73)[44]:

(73) NONFINALITY (1)
|
ALIGN-R (1)
*[+ SPREADGLOTTIS]/LOW (·524)
|
*H

The probability ranking of the constraint *[+ SPREADGLOTTIS]/LOW over *H, ·524, is crucial to account for the accent variation driven by aspiration in the novel forms. For example, this

[44] One set of ranking values that yield the probabilities presented in this section is: NONFINALITY 106.00;H]$_{PW}$ 104.00; ALIGN-R 90.00; *[+SPREADGLOTTIS]/LOW -2,675.85; *H -2,658.01.

ranking predicts the output with double accent at a ·524 probability and the output with penultimate accent at a ·476 probability, given an input like a novel form /t^hakapa/. The stochastic evaluation over a novel form /t^hakapa/ is illustrated as follows:

(74) Aspiration-driven variations

Input ~ Output	NONFINAL (1)	ALIGN-R (1)	*[+S.C]/L (0.524)	*H
☞ a. |t^ha.ka.pa| ~ /t^há.ká.pa/ (0.524)		*		**
☞ b. |t^ha.ka.pa| ~ /t^ha.ká.pa/ (0.476)		*	*	*
c. |t^ha.ka.pa| ~ t^há.ka.pa		**!		*
d. |t^ha.ka.pa| ~ t^ha.ka.pá	*!		*	*
e. |ta.ka.pa| ~ /tá.ká.pa/		*		**!
☞ f. |ta.ka.pa| ~ /ta.ká.pa/ (1)		*		*

The candidates with initial accent and with final accent (74c) and (74d) are ruled out by the constraints NONFINALITY and ALIGN-R, respectively. The candidate with either double accent (74a) or penultimate accent (74b) will be selected depending on the ranking of the constraints *[+SPREADGLOTTIS]/LOW and *H. Under the probability ranking ·524 of *[+ SPREADGLOTTIS]/LOW >> *H, the double accented candidate (74a) wins over the penultimate accented candidate (74b) at a ·524 probability, whereas the penultimate accented candidate (74b) will be selected at a ·476 probability. In contrast, only the candidate with penultimate accent (74f) will be the winner when the input does not contain aspirated onset consonants.

The aspiration effects on accent patterns emerging in novel words manifest a hidden constraint *[+SPREADGLOTTIS]/LOW, which does not operate actively in native accent placement. The emergence of the constraint *[+SPREADGLOTTIS]/LOW instantiates the emergence of Universal Grammar, since NKK native accent patterns do not clearly substantiate this constraint. Furthermore, stochastic ranking of the relevant constraints explains the tendencies toward high toned syllables with aspirated onset consonants.

7.5.3 Alternative account: dialectal influence?

The effect of aspiration on the F0 of a following vowel has been considered as one of the features signaling the phonation contrast among lax, tense and aspirated stops in standard Korean. Voice onset time (VOT) has also been assumed to be a robust phonetic cue for the laryngeal contrast. However, Silva (2006) claimed that the phonetic implementation of the contrast has changed over the past two generations: the VOT distinction between lax and aspirated stops has been neutralized for some younger speakers and F0 has been utilized as a primary cue for the distinction between lax and aspirated stops. Silva argued that contemporary Seoul Korean, which is not a pitch accent/tone language, has recently developed a tonal system to encode the phonemic contrast between lax and aspirated stops: lax stops are associated with a low tone while aspirated stops are marked with a high tone. Silva's argument is supported by a perception study (M-R Kim *et al.* 2002), which showed that stop identification was strongly determined by information on the vowel (F0) rather than on the consonant (VOT), with low F0 cueing lax stops and higher F0 cueing aspirated stops. Therefore, for the two-way contrast between lax stops and aspirated stops, vowel F0 is sufficient to cue the distinction. In addition, H-K Jun (2006) also found that aspiration was a relevant factor in accent patterns of loanwords in South Kyungsang Korean (SKK): initial aspirated consonants induced high tones in SKK loans (eg., *tʰótʰə́l* 'total'; *cʰénə́l* 'channel'; *kʰémpʰə́sɨ* 'campus'). Considering that the onset-sensitive NKK accentuation of novel words was in accordance with the changed phonetic implementation of the stop contrast in contemporary standard (or Seoul) Korean or with the association of high tones and aspiration in SKK, it is plausible that NKK speakers would perceive the tonal effect of aspiration as a cue for lexical accent when they encounter new words spoken in standard Korean or other dialects and apply it in NKK accent placement.

7.6 Conclusion

Aspirated onset consonants are not necessarily associated with a high tone pitch accent in North Kyungsang Korean although vowels following aspirated obstruent consonants tend to have higher F0 than those following other consonants. However, the current study presented evidence that aspiration plays a role in the assignment of a high tone pitch accent to novel words. Although penultimate accent was preferred both in bisyllabic words and in trisyllabic words, the predominant penultimate accentuation was disturbed in trisyllabic words when a word began with aspirated onset consonants. As for bisyllabic words, final accent was more frequent when final syllables contained an aspirated onset consonant.

This study suggests that even though the effect of aspiration on accentuation was not categorical, the phonetic effect was important enough to serve as a factor at least in deriving variation in accent patterns in newly adopted words. The results suggest that for words without underlying accent, penultimate accent is optimal only when a word does not contain aspirated consonants. Double accent or final accent are variant forms for some cases when a word contains an aspirated consonant in the initial syllable or in the final syllable. The association of aspiration with a high tone is attributable to a constraint, *[+SPREADGLOTTIS]/LOW. Namely, some speakers who are more sensitive to the association between aspiration and higher F0 would have a grammar which ranks the constraint *[+SPREADGLOTTIS]/LOW higher than other constraints, e.g. *H. On the other hand, other speakers who are not sensitive to the phonetic effects would have a grammar with the constraint *[+SPREADGLOTTIS]/LOW ranked low, which will produce dominant penultimate accentuation.

Appendix IV: Results of Experiment 3

Table 22 Results in bisyllabic words in Experiment 3

Word Type	Novel word	Accent in novel words			Total
		Penult	Final	Double	
NN	tapa	25	1	4	30
	kapa	26	0	4	30
	piki	22	2	6	30
	cipi	24	0	6	30
AN	t^hapa	27	0	3	30
	k^hapa	24	1	5	30
	p^hiki	29	0	1	30
	c^hipi	26	0	4	30
NA	tap^ha	21	5	4	30
	kap^ha	19	6	5	30
	pik^hi	25	1	4	30
	cip^hi	22	4	4	30
AA	t^hap^ha	24	1	5	30
	k^hap^ha	25	1	4	30
	p^hik^hi	22	3	5	30
	c^hip^hi	23	2	5	30
Total		384	27	69	480

Table 23 Results in trisyllabic wordsin Experiment 3

Word Type	Novel word	Accent in novel words			Total
		Penult	Final	Double	
NNN	takapa	21	2	7	30
	pakapa	26	1	2	29
	kitici	21	1	8	30
	citiki	22	1	7	30
ANN	t^hakapa	16	1	13	30
	p^hakapa	18	1	11	30
	k^hitici	7		23	30
	c^hitiki	14	1	15	30
NAN	tak^{h}apa	27	1	1	29
	pak^{h}apa	20	1	9	30
	kit^{h}ici	19	2	7	28
	cit^{h}iki	24	1	5	30
AAN	t^hak^{h}apa	20	1	7	28
	p^hak^{h}apa	19	1	9	29
	k^hit^{h}ici	12	1	16	29
	c^hit^{h}iki	19	1	9	29
Total		305	17	149	471

Chapter 8 Conclusion and Future Directions

This dissertation investigated factors which determine NKK accent patterns in novel forms. Following previous findings, four factors were investigated: syllable structure (e.g. Kenstowicz and Sohn 2001; Y-H Chung 2002, 2006; Shinohara 1997, 2000; Guion *et al*. 2003; Face 2004); lexical frequency (e.g. Zuraw 2000; Frisch & Zawaydeh 2001; Hayes and Londe 2006); analogy to patterns in phonetically similar words (e.g. Eddington 2000, 2004; Face 2004; Albright 2008); and aspiration effects of syllable onsets (e.g. S-A Jun 1993; Silva 1992, 2006; M-R Kim 2000; M-R Kim and Duanmu 2004; Kenstowicz and Park 2006).

The findings in this dissertation provide insight into a long-standing puzzle in Korean linguistics: the fact that while the accent pattern in native words must for the most part be lexically specified, the accent patterns in loanwords seem to be at least partialy predictable from the syllable structure of the word. I have argued that the syllable-sensitive patterns found in loanwords have their source both in the statistical tendencies of the lexicon and in universal tendencies, which favor the enforcement of universal markedness constraints in the absence of lexical specification. The stochastic grammar emerging from the native vocabulary favors syllable-sensitive patterning. Finally, I suggested that this proposed stochastic grammar reflects the final state of the grammar. The initial ranking among hidden markedness constraints NONFINALITY, ALIGN-R, *C_{μ} (no moraic codas) and WEIGHT-BY-POSITION (moraic codas) does not remain fixed at the final state, but rather must be adjusted in accordance with the lexical frequency patterns throughout the learning process. That is, NKK learners re-rank constraints gradually as they encounter counterexamples in the lexicon. Therefore, the markedness constraints at the final state end up being ranked stochastically and with sufficient overlap to generate variation. I propose that re-ranking of the constraints is not necessary in loanword phonology and that the uniform ranking of the markedness constraints as in the native phonology can also explain the patterns emerging in loanwords.

Although syllable structure does not categorically predict NKK accent patterns in native words, experimental studies revealed that NKK speakers accent novel words in a patterned manner in accordance with syllable structure. Namely, CVC syllables tend to attract accent; otherwise, penultimate accent is preferred. The syllable structure-sensitive tendencies in novel words are actually consistent with the regular patterns emerging in loanwords. Kenstowicz and Sohn (2001) suggest that the regular patterns in loanwords reflect the "UG default status" of closed syllables as heavy. A study of the lexical statistics of accent patterning in NKK showed that the structure-sensitive accent tendencies in novel words were generally consistent with the statistical accent patterns in the lexicon, though the effect of syllable structure was even stronger in novel words. This suggests that the native lexicon does provide evidence for this UG setting so that NKK speakers can learn the regular patterns on the basis of the native accent patterns in existing words and apply them to newly adopted words, where there is no pressure to preserve lexically specified accent patterns.

I have also argued that the accentuation of novel words cannot be explained solely as an effect of analogy to phonetically similar words. Such an effect would be predicted in Exemplar Theory models (e.g. Bybee 1985, 2001, 2006; Pierrehumbert 2001a), which assume that lexical items are stored in a network-like multi-dimensional organization and that the recognition process involves activation of a cloud of exemplars. Thus, this theory predicts that production of novel words should draw on activation of neighboring lexical items which are similar to the target novel item. However, the study of NKK accent patterns using novel words with phonetically-similar existing counterparts demonstrated that the effect of syllable structure was stronger than the analogy effect: NKK speakers accented novel words with obvious existing counterparts in accordance with the syllable structure-sensitive tendencies, even when accent patterns in analogous existing words contradicted the syllable structure-sensitive tendencies. Therefore, I conclude that the effects of syllable structure emerging in novel words are indeed attributable to a phonologically-based source rather than an analogically-based process.

Additionally, I have shown that vowel quality was another factor which influenced accent assignment in novel words: word-initial high vowels were more likely to attract accent than word-initial low vowels. This high vowel effect could be attributable to the inherent phonetic quality of high vowels: high vowels have intrinsically higher pitch than lower vowels. This suggests that phonetic qualities (e.g. F0) might also contribute to NKK accent placement, leading speakers to more readily associate this intrinsic high pitch with the accentual high tone. Furthermore, aspiration was seen to play a role in accentuation of novel words: a vowel following an aspirated onset tended to attract a high tone pitch accent. Although penultimate accent was most frequent in bisyllabic and trisyllabic words containing only light syllables, non-penultimate accent was more likely in words containing aspirated onset consonants in word-initial (in trisyllabic words) or word-final syllables (in bisyllabic words). All these results imply that phonetic factors as well as phonologically-grounded constraints play a role in NKK accentuation of new words.

However, all the patterns emerging in novel words represented statistical tendencies rather than categorical generalizations. I proposed that a stochastic grammar provides an account of the probabilistic accent patterns in novel words. Given the fact that the structure-sensitive tendencies in novel words were consistent with the lexical statistical patterning, I argue that NKK speakers internalize a grammar that reflects the lexical statistical patterning, and this grammar projects the patterning onto novel forms. Therefore, NKK speakers construct a stochastic grammar which is equipped with probabilistic constraint ranking on the basis of the native lexical accent patterning. Crucially, however, the comparison of the corpus study and the accentuation of novel words revealed that the structure-sensitive tendencies emerged more strongly in novel words than in the lexicon. This implies that NKK learners do not simply replicate the lexical frequency of different accent patterns, but rather are biased toward the structure-sensitive patterns, even though a number of existing words have lexically specified accent patterns that contradict the syllable-sensitive pattern.

These results cast light on the question of the general relationship between statistical tendencies in the lexicon and speakers' behavior with respect to novel forms, and specifically on the question of what sorts of statistical tendencies are likely to productively extend to new forms. Becker *et al*. (2008) argue that only phonologically-grounded generalizations show up in novel forms. Their study involves Turkish nouns, in which some final stops alternate in voicing, while others do not. These voicing alternations are not predictable, but the probability of alternation correlates with specific independent properties: the size of the noun (polysyllabic nouns are more likely to show voicing alternations than monosyllables); the place of articulation of the final stop (coronals are more likely to alternate than non-coronals); and the quality of the preceding vowel (alternation is more likely following a high vowel). Cross-linguistically, voicing alternations are associated with the first two of these factors, word size and consonant place, but not with the third factor, vowel height, suggesting that this association in Turkish is accidental rather than phonologically-based. Becker *et al.* found that Turkish speakers' production of nonce words does not show a correlation with vowel height, but does show an effect of word size and consonant place. They propose that Universal Grammar filters out accidental generalizations, so that Turkish speakers project only phonologically-grounded lexical trends onto novel words.

Since the present study revealed several statistical tendencies in the lexicon (penultimate accent in words without heavy syllables; final accent in words ending with a heavy syllable; and double accent in words beginning with a heavy syllable), the next question to ask is whether the tendencies that are extended to novel forms are those that represent common cross-linguistic patterns. Cross-linguistically, the connection between prominence and syllable weight is fairly common: for example, heavy syllables tend to attract stress. The preference for stressing the heavy syllable has been described with a constraint PeakProminence (Prince & Smolensky 1993), which favors assignment of prominence to the heavy syllable. De Lacy (2002) also found some connection between high tone and stress in the tone-sensitive stress patterns in Ayutla, a Mixtec language: higher toned syllables are preferred as heads of feet over lower-toned ones. He proposes that higher tone is more prominent than lower tone, suggesting a universal Tone Prominence scale, H > M > L. On this view, the connection between high tone and heavy syllables in NKK reflects universal tendencies concerning the relationship between syllable weight and prominence as expressed by stress, accent, or high tone. Thus, I propose that in NKK accentuation of novel words, speakers show a bias toward the grammatically-motivated patterns, as in Turkish. In accenting novel words, speakers do not simply mirror the statistics of the lexicon; instead they favor patterns that reflect universal preferences.

Nonetheless, the CVC heavy criterion for a high tone pitch accent which emerged in NKK novel words still remains puzzling, considering cross-linguistic tendencies. This CVC heavy weight criterion is rare for tone cross-linguistically, although it is common for stress. According to Gordon (1999; 2004), CVC syllables with obstruent codas do not count as heavy for tone in most of the languages which display weight-sensitive tone patterns; e.g., contour tones are allowed only on CVV or CVC[+son] syllables in many languages. Gordon's cross-linguistic survey showed that CVV and CVC[+son] syllables count as heavy in 25 out of the 49 languages having a weight-sensitive tone system, and that CVV syllables but not CVC[+son] syllables count as heavy in 21 out of the 49 languages. Only 3 languages among the languages observe the CVC heavy weight criterion for tone. Recall that CVC[+son] syllables did not attract accent consistently more often than CVC[-son] in NKK (presented in Chapter 3), which suggests that NKK does not necessarily observe the CVV, CVC[+son] heavy weight criterion. Gordon (1999; 2004) proposed a phonetically-based explanation of the different syllable weight criteria in different languages. According to Gordon, languages adopt weight distinctions which are most effective or sensible phonetically: namely, weight distinctions rely on groups of syllables that are maximally differentiated based on the phonetic property of total energy. For example, in Khalka, where CVV syllables are treated as heavy but not CVC syllables, the phonetic distinction in total energy was most visible between VV and other rhymes, while in Finnish, where both CVV and CVC syllables are treated as heavy, the phonetic distinction between VV/VC and other rimes was most visible. Although this dissertation provided phonological accounts of the connection between CVC syllables and high tone in Chapter 6, it would be interesting if the emergence of the CVC heavy criterion could also have some support from NKK phonetics.

One study for future research would investigate the possible phonetic basis of the attraction of pitch accent to heavy syllables in NKK. Under Gordon's view, CVC syllables as well as CVV syllables should be phonetically distinct from CV syllables in NKK in a way that could lead NKK speakers to associate those syllables with a high tone. Therefore, studies of the acoustic correlates of syllable weight in NKK will be necessary to validate this phonetically-based approach and shed light on additional sources of the emergence of heavy CVC syllables in NKK.

Alternatively, the association between high tone and CVC syllables could be a genuine reflection of universal preferences, independent of the language-specific phonetics. Experimental studies which involve the learning of artificial languages will provide a good test ground to verify this hypothesis on the universal association between high tone and CVC syllables. If accented CVC syllables indeed reflect universal preferences, non-native NKK speakers who are not speakers of pitch accent languages (e.g. a language with a weight "insensitive" prosody system) should also show a bias toward accented closed syllables in learning artificial languages, given accented CVC syllables and accented CV syllables with equal frequency.

Another puzzling question comes from the role of intrinsic effects in tone production. This dissertation demonstrated that NKK speakers were more likely to associate intrinsic high pitch with the accentual high tone: word-initial high vowels were more likely to attract a high tone than word-initial low vowels; and vowels following an aspirated consonant tended to attract a high tone. The intrinsic effects of preceding consonants in tone production have been commonly found in the historical development of tones: high and low tones have developed after initial voiceless and voiced consonants respectively in many languages (e.g. Hombert 1978; Hombert *et al.* 1979). Several scholars (Hombert 1978; Ohala 1993) suggested a perceptual explanation for the tonal development in that listeners perform a 'dissociation' parsing error on F0 from the consonant and reinterpret a previously intrinsic cue (F0) as a tonal distinction. This explanation could apply to the association between aspiration and high tone emerging in NKK accent patterns, which has been also found in several other languages (e.g. S-A Jun 1993; S-H Lee 2008). A perception study needs to be done to substantiate this perceptual explanation of the aspiration effects in NKK accent patterns.

On the other hand, intrinsic effects of vowel height on tonal development have not been widely attested cross-linguistically (Hombert 1978; Hombert *et al.* 1979). Hombert *et al.* (1979) conjectured that the lack of vowel height effects in the development of tones might be because intrinsic effects caused by preceding consonants and by vowel height could be perceived differently: namely, the F0 perturbation by consonants may be more noticeable to listeners because it involves dynamic changes in F0, while the effects of vowel quality are realized with different F0 steady-state levels. However, the vowel height effects which showed up in NKK accent patterns contradict this cross-linguistic tendency. This raises several questions: (i) where did the accent patterns sensitive to vowel height come from? (ii) was the vowel height effect related to language-specific phonetics, or (iii) could the effect be ascribed to lexical frequency? It would be of interest if lexical statistics of accent patterns in existing native words support the vowel effects in NKK accentuation. If so, this suggests that the vowel quality-sensitive patterns should have a basis in language-particular patterns. In contrast, the intrinsic effects of aspiration might not necessarily have lexical support from existing words, which then clearly demonstrates that the aspiration effects should have a basis in universal tendencies. Thus, a study of the statistical association in the lexicon between tone and vowel quality vs. tone and aspiration is suggested to clarify this remaining issue.

Furthermore, recall that in Chapter 6, a simulation assuming the Gradual Learning Algorithm (GLA)(Boersma1997; Boersma & Hayes 2001) predicted double accent with a lower frequency than the actual number of double accents found in novel words. I assumed that this mismatch was because the simulation ignored the phonetic effects (Section 6.4.1;6.4.2). It would be worth testing whether adopting constraints for the phonetic effects in the simulation will compensate for the mismatch of double accent assignment.

To conclude, the findings in this dissertation shed light on the question of what determines phonological behavior. This dissertation has demonstrated that NKK speakers' behavior in accenting novel words reflects an abstract internalized grammar rather than analogical inference. Lexical frequency plays a role in shaping the grammar, inducing variation and probabilistic preferences, which suggests that the grammar is not necessarily categorical. Furthermore, the effects of lexical frequency interact with grammartical effects in that patterns that are consistent with universal preferences are more likely to extend to novel words.

REFERENCES

Albright, A. 2008. Modeling analogy as probabilistic grammar. Ms. MIT. (Downloadable from http://web.mit.edu/albright/www/papers/Albright-ModelingAnalogyAsGrammar.pdf)

Albright, A. & Hayes, B. 2002. Modeling English past tense intuitions with minimal generalization. SIGPHON 6: *Proceedings of the Sixth Meeting of the ACL Special Interest Group in Computational Phonology*. 58-69.

Albright, A. & Hayes, B. 2003. Rules vs. analogy in English past tenses: A computational/experimental Study. *Cognition* 90. 119-161.

Allopenna, P., Magnuson, J., & Tanehaus, M. 1998. Tracking the time course of spoken word recognition using eye movements: Evidence for continuous mapping models. *Journal of Memory and Language* 38. 419-439.

Anttila, A. 1997. Deriving variation from grammar. In Hinskens, F., van Hout, R., & Wetzels, L. (eds.), *Variation, Change and Phonological Theory*. 35-68. Amsterdam: John Benjamins.

Apoussidou, D. 2007. *The Learnability of Metrical Phonology*. PhD dissertation, University of Amsterdam.

Baker, R. & Smith, P. 1976. A psycholinguistic study of English stress assignment rules. *Language and Speech* 19. 9-27.

Becker, M., Ketrez, N., & Nevins, A. 2008. The surfeit of the stimulus: Analytic biases filter lexical statistics in Turkish devoicing neutralization. Ms. Reed University, Yale University, and Harvard University. (Downloadable from http://becker.phonologist.org/papers/becker_ketrez_nevins_turkish_voicing.pdf)

Boersma, P. 1997. How we can learn variation, optionality, and probability. Ms. University of Amsterdam.

Boersma, P. 1998. *Functional Phonology: Formalizing the Interaction between Articulatory and Perceptual Drives*. PhD dissertation, University of Amsterdam.

Boersma, P. 2001. Phonology-semantics interaction in OT, and its acquisition. In Kirchner, R., Wikeley, W., & Pater, J. (eds.), *Papers in Experimental and Theoretical Linguistics* 6. 4-35. Edmonton: University of Alberta. [ROA 369, 1999].

Boersma, P. & Hayes, B. 2001. Empirical tests of the Gradual Learning Algorithm. *Linguistic Inquiry* 32. 45-86.

Boersma, P. & Weenink, D. 2003. Praat: Doing phonetics by computer (version 4.3.17). [Computer program]. Retrieved in 2003 from <http://www.praat.org/>

Broselow, E., Chen, S., & Huffman, M. 1997. Syllable weight: Convergence of phonology and phonetics. *Phonology* 14. 47-82.

Broselow, E., Chen, S., & Wang, C. 1998. The emergence of the unmarked in second language phonology. *Studies in Second Language Acquisition* 20. 261-280.

Bybee, J. 1985. *Morphology: A study of the Relation between Meaning and Form*. Amsterdam: John Benjamins.

Bybee, J. 2000. Lexicalization of sound change and alternating environments. In Broe, M. & Pierrehumbert, J. (eds.), *Laboratory Phonology V: Acquisition and the Lexicon*. 250-268. Cambridge: Cambridge University Press.

Bybee, J. 2001. *Phonology and Language Use*. Cambridge, U.K.: Cambridge University Press.

Bybee, J. 2006. From usage to grammar: The mind's response to repetition. *Language* 82, 711-733.

Bybee, J. & Hopper, P. 2001. *Frequency and the Emergence of Linguistic Structure*. Amsterdam: John Benjamins.

Chung, K. 1980. *Neutralization in Korean: A Functional View*. PhD dissertation, University of Texas at Austin.

Chung, Y-H. 1991. *The Lexical Tone System of North Kyungsang Korean*. PhD dissertation, Ohio State University.

Chung, Y-H. 2002. Contextually-dependent weight of closed syllable of the North Kyungsang dialect and English, *Eoneohag* 33. 23-35. The Linguistic Society of Korea.

Chung, Y-H. 2006. Emergence of universal grammar in loanword phonology of North Kyungsang Korean. *Eoneohag* 44. 3-35. The Linguistic Society of Korea.

Coleman, J. & Pierrehumbert, J. 1997. Stochastic phonological grammars and acceptability. *Proceedings of the Third Meeting of the ACL Special Interest Group in Computational Phonology*. 49-56. Association for Computational Linguistics, Somerset, NJ.

Crosswhite, K., Alderete, J., Beasley, T., & Markman, V. 2003. Morphologically effects on default stress in novel Russian words. In Garding, G. & Tsujimura, M. (eds.), *WCCFL 22 Proceedings.* 151-164. Somerville, MA: Cascadilla Press.

Crothers, J. 1978. Typology and universals of vowel systems. In Greeberg, J. (ed.), *Universals of Human Language, Volume 2, Phonology*. 93-152. Stanford, CA: Standford University Press.

Daelemans, W., Gillis, S., & Durieux, G. 1994. The acquisition of stress: A data-oriented approach. *Computational Linguistics* 20. 421-451.

Daelemans, W., Zavrel, J., Van der Sloot, K., & Van den Bosch, A. 2000. TiMBL: Tilburg memory based learner reference guide 3.0. Report 00-01, Computational Linguistics Tilburg University.

Davidson, L., Jusczyk. P., & Smolensky, P. 2004. Implication and exploration of Richness of the Base. In Kager, R., Pater, J., & Zonneveld, W. (eds.), *Constraints in Phonological Acquisition.* 321-368. Cambridge: Cambridge University Press.

de Lacy, P. 2002. *The Formal Expression of Markedness*. PhD dissertation. University of Massachusetts at Amherst. Mass.: GLSA Publications.

Demuth, K. 1995. Markedness and the development of prosodic structure. *NELS*. 13-25. Amherst, Mass.: GLSA, University of Massachusetts.

Eddington, D. 2000. Spanish stress assignment within the Analogical Modeling of Language. *Language* 76. 92-109.

Ernestus, M. & Baayen, H. 2003. Predicting the unpredictable: interpreting neutralized segments in Dutch. *Language* 79. 5-38.

Face, T. 2004. Perceiving what isn't there: Non-acoustic cues for perceiving Spanish stress. In Face, T. (ed.), *Laboratory Approaches to Spanish Phonology*. 117-141. Berlin: Mouton de Gruyter.

Frisch, S. & Zawaydeh, B. 2001. The psychological reality of OCP-Place in Arabic. *Language* 77. 91-106.

Gillis, S., Daelemans, W., & Durieux, G. 2000. "Lazy Learning": A comparison of natural and machine learning of stress. In Broeder, P. & Murre, J. (eds.), *Models of Language Acquisition*, 76-99. Oxford: Oxford University Press.

Gim, C-G. 1977. *Kyungsangdo Pangeonui Sengco Chegye* (Tone System of Kyungsang Korean). PhD dissertation, Seoul National University.

Gim, C-G. 1985. A way to describe the tonal correspondence between Middle Korean and Kyungsang Dialects. *Historical Linguistics*. 203-240. Seoul: Cenyeywen.

Gim, C-G. 1994. 60tae Daegu pangeon kwa 20tae Daegu pangeon seongco ui taeco punseok (The comparison of tone patterns in Daegu dialect between sixties and twenties). *Eoneohak* 13. 23-80.

Gim, C-G. 2003. *Youngnam Pangeon Seongco Pigyo* (The comparison of tone patterns in Youngnam dialects). Seoul: Yuklak.

Gnanadesikan, A. 2004. Markedness and faithfulness in child phonology. In Kager, R., Pater, J., & Zonneveld, W. (eds.), *Constraints in Phonological Acquisition*. 73-108. Cambridge: Cambridge University Press.

Gordon, M. 1999. *Syllable Weight: Phonetics, Phonology, and Typology*. PhD dissertation, University of California, Los Angeles.

Gordon, M. 2004. Syllable weight. In Hayes, B., Kirchner, R., & Steriade, D. (eds.), *Phonetically Based Phonology*. 277-312. New York: Cambridge University Press.

Gordon, M. 2005. A perceptually-driven account of onset-sensitive stress. Ms. University of California, Santa Barbara.

Guion, S., Clark, J., Harada, T. & Wayland, R. 2003. Factors affecting stress placement for English nonwords include syllabic structure, lexical class, and stress patterns of phonologically similar words. *Language and Speech* 46:4. 403-427.

Han, M. & Weitzman, R. 1965. *Studies in the Phonology of Asian Languages III: Acoustic Characteristics of Korean Stop Consonants*. Los Angeles: Acoustic Phonetics Research Laboratory, University of Southern California.

Han, M-S. 1964. Duration of Korean vowels. Technical Report No. 2, University of California, Los Angeles.

Hayes, B. 1989. Compensatory lengthening in moraic phonology. *Linguistic Inquiry* 20. 253-306.

Hayes, B. 1995. *Metrical Stress Theory: Principles and Case Studies*. Chicago: University of Chicago Press.

Hayes, B. 2004. Phonological acquisition in Optimality Theory. In Kager, R., Pater, J., & Zonneveld, W. (eds.), *Constraints in Phonological Acquisition*. 158-203. Cambridge: Cambridge University Press.

Hayes, B. & Londe, Z. 2006. Stochastic phonological knowledge: The case of Hungarian vowel harmony. *Phonology* 23. 59-104.

Hayes, B. & Wilson, C. 2008. A Maximum Entropy Model of phonotactics and phonotactic learning. *Linguistic Inquiry* 39:3. 379-440.

Hay, J., Pierrehumbert, J., & Beckman, M. 2004. Speech perception, well-formedness and the statistics of the lexicon. In Local, J., Ogden, R. & Temple, R. (eds.), *Phonetic Interpretation: Papers in Laboratory Phonology VI*. 58-74. Cambridge: Cambridge University Press.

Hombert, J-M. 1978. Consonant type, vowel quality and tone. In Fromkin, V. (ed.), *Tone: A Linguistic Survey*. 77- 111. New York: Academic Press.

Hombert, J-M., Ohala, J. & Ewan, W. 1979. Phonetic explanation for the development of tones.

Language 55:1. 37-58. Linguistic Society of America.

Hoole, P., Kühnert, B., and Pouplier, M. 2012. System-related variation. In Cohn, A., Fougeron, C., & Huffman, M. (eds.), *The Oxford Handbook of Laboratory Phonology*. 115-132. New York: Oxford University Press.

Hsieh, F. & Kenstowicz, M. 2008. Phonetic knowledge in tonal adaptation: Mandarin and English loanword in Lhasa Tibetan. *Journal of East Asian Linguistics* 17. 279-297.

Huh, W. 1965. *Kuk-eo-Eum-un-hak* (Korean Phonology). Seoul: Chungeumsa.

Huh, W. 1985. Tone in Kyungsang dialects. *Korean Journal* 25. 19-32.

Hyman, L. 1985. *A Theory of Phonological Weight*. Dordrecht: Foris.

Idsardi, W. & Chang, W-H. 2003. Markedness and the phonetic interpretation of tone in North Kyungsang Korean. In Nowak, P., Yoquelet, C., & Mortensen, D. (eds.), *Proceedings of the 29th annual meeting of the Berkeley Linguistics Society*. 197-208. Berkeley, CA: Berkeley Lingusitcs Society.

Jun, H-K. 2006. Factors affecting accentual patterns of loanwords in Pusan Korean. In Kuno, Susumu *et al*. (eds.), *Harvard Studies in Korean Linguistics* 11. 158-170.

Jun, J-H., Kim, J-S., & Jun, S-A. 2006. The prosodic structure and pitch accent of Northern Kyungsang Korean. *Journal of East Asian Linguistics* 15. 289-317.

Jun, S-A. 1993. *The Phonetics and Phonology of Korean Prosody*. PhD dissertation, Ohio State University.

Jusczyk, P., Luce, P., & Charles-Luce, J. 1994. Infants' sensitivity to phonotactic patterns in the native language. *Journal of Memory and Language, 33*. 630-645.

Kenstowicz, M. & Sohn, H. 2001. Accentual adaptation in North Kyungsang Korean. In Kenstowicz, M. (ed.), *Ken Hale: A Life in Language*. 239-270. Cambridge, MA: MIT Press.

Kenstowicz, M. & Park, C. 2006. Laryngeal features and tone in Kyungsang Korean: A phonetic study. *Studies in Phonetics, Phonology, and Morphology*. 12.2. 247-264.

Kenstowicz, M., Cho, H., & Kim, J. 2007. Contrasts, mergers, and acquisition in Kyungsang accent. Ms. MIT. (Downloadable from http://web.mit.edu/linguistics/people/faculty/kenstowicz/publications.html)

Kenstowicz, M. 2011. Accent classes and lexical drift in Kyungsang Korean. Handout for MIT Phonology Circle. Oct. 31, 2011.

Kim, C-W. 1965. On the autonomy of the tensity feature in stop classification (with special reference to Korean stops). *Word* 21. 339-359.

Kim, C-W. 1968. Review of Lieberman 1967. *Language* 44. 830-842.

Kim, G-R. 1988. *The Pitch-accent System of the Taegu Dialect of Korean with Emphasis on Tone Sandhi at the Phrase Level*. PhD dissertation. University of Hawaii.

Kim, H-J. 2003. *Interaction between Tone in North Kyungsang Korean and Stress in English.* MA thesis. Kyungpook National University.

Kim, H-J. 2011. The phonetics and phonology of onset-sensitive accentuation of novel words in North Kyungsang Korean. *Studies in Phonetics, Phonology, and Morphology* 17:2. 193-214.

Kim, H-J. In press. Emergent hidden grammar: Stochastic patterning in Korean accentuation of novel words. *The Proceedings of the 36th Annual Meeting of the Berkely Linguistics Society*. eLanguage.

Kim, H-J. In press. Analogy effects on North Kyungsang Korean accentuation of novel words. In Nam, S-H., Ko, H-J., & Jun, J-H. (eds.), *Japanese/Korean Linguistics 21*. Stanford, CA.: CSLI.

Kim, J-W. 1991. Kyungsang pangeonui ponjilkwa Jungse Kukeoui seongjo taeunge taehayeo (About the essense of Kyungsang dialect and response to the Middle Korean tone). *Eoneohag* 13. 75-93.

Kim, M-R. 2000. An alternative account of so-called lax consonants with special reference to consonant-tone interaction. *Studies in Phonetics, Phonology, and Morphology* 6. 333-352.

Kim, M-R, Beddor, P. S., & Horrocks, J. 2002. The contribution of consonantal and vocalic information to the perception of Korean initial stops. *Journal of Phonetics* 30. 77-100.

Kim, M-R, & Duanmu, S. 2004. 'Tense' and 'lax' stops in Korean. *Journal of East Asian Linguistics* 13. 59-104.

Kim, N-J. 1997. *Tone, Segments, and Their interaction in North Kyungsang Korean: A Correspondence Theoretic Account*. PhD dissertation, Ohio State University.

Kim, S-H. 1999. *The Metrical Computation in Tone Assignment*. PhD Dissertation, University of Delaware.

Kubozono, H. 2006. Where does loanword prosody come from? A case study of Japanese loanword accent. *Lingua* 116. 1140-1170.

Ladd, D. R. & Silverman, K. 1984. Vowel intrinsic pitch in connected speech. *Phonetica* 41. 31-40.

Lee, C-H. 1955. *Practical Korean Grammar*. Seattle: University of Washington Press.

Lee, H-Y. 1994. Seoulmalkwa Kyungsangdo pangeonui unyul yuhyungron (The prosodic typology of Seoul and Kyungsang Korean). *Eoneohag* 15. 353-382.

Lee, H-Y. 1997. Kukeo unyulron (Korean prosody). *Korean Studies Series,* V 65. Seoul: The Korean Research Center.

Lee, D-M. 2008a. The tonal structures and the locations of the main accent of Kyungsang Korean words. *IULC Working Papers.* (Downloadable from https://www.indiana.edu/~iulcwp/pdfs/08-Lee.pdf)

Lee, D-M. 2008b. *The Loanword Tonology of South Kyungsang Korean*. PhD dissertation. Indiana University.

Lee, H-J. & Jongman, A. 2010. Effects of tone on the three-way laryngeal distinction in Korean: An acoustic and aerodynamic comparison of the Seoul and South Kyungsang dialects. Paper presented at *The 159th Meeting of the Acoustical Society of America*, Baltimore, Maryland.

Lee, S-H. 2008. *Consonant-Tone Interaction in Optimality Theory*. PhD Dissertation, Rutgers, The State University of New Jersey, New Brunswick.

Levelt, C. 1995. Unfaithful kids: place of articulation patterns in early child language. Paper presented at the Department of Cognitive Science, Johns Hopkins University. Baltimore, MD.

Levelt, C. & van de Vijver, R. 2004. Syllable types in cross-linguistic and developmental grammars. In Kager, R., Pater, J., & Zonneveld, W. (eds.), *Constraints in Phonological Acquisition*. 204-218. Cambridge: Cambridge University Press.

Lisker, L, & Abramson, A. 1964. A cross-language study of voicing in initial stops: Acoustical measurements. *Word 20*. 384-422.

McCarthy, J. 1979. *Formal Problems in Semitic Phonology and Morphology*. PhD. Dissertation, MIT, Cambridge, MA.

McCarthy, J. & Prince, A. 1994. The emergence of the unmarked: optimality in prosodic morphology. In Gonzàlez, M. (ed.), *Proceedings of the NELS* 24. 333-379. Amherst: GLSA.

McCarthy, J. & Prince, A. 1995. Faithfulness and reduplicative identity. *University of Massachusetts Occasional Paper* 18. [ROA-60, http://ruccs.rutgers.edu/roa.html]

Magen, H.S. & Blumstein, S.E. 1993. Effects of speaking rate on the vowel distinction in Korean. *Journal of Phonetics 21*. 387-409.

Marslen-Wilson, W. 1987. Functional parallelism in spoken word recognition. *Cognition 25*. 71-102.

Narahara, T. 1985. The accentual system of the Korean Kyungsang Dialect. In Kuno, S. *et al*. (eds.), *Harvard Studies in Korean Linguistics I*. 257-69. Seoul: Hanshin Publishing Company.

Nosofsky, R.M. 1990. Relations between exemplar similarity and likelihood models of classification. *Journal of Mathematical Psychology 34*. 393-418.

Ohala, J. 1993. Sound change as nature's speech perception experiment. *Speech Communication* 13. 155-161.

Park, J-K. 1985. A phonetic and phonological study of duration in modern Korean. MA thesis. Seoul National University.

Pater, J. & Paradis, J. 1996. Truncation without templates in child phonology. *Proceedings of the Boston University Conference on Language Development.* 540-552. Somerville, Mass.: Cascadilla Press.

Pierrehumbert, J. 2001a. Exemplar dynamics: Word frequency, lenition and contrast. In Bybee, J. & Hopper, P. (eds.), *Frequency and the Emergence of Linguistic Structure.* 137-158. Amsterdam: John Benjamins.

Pierrehumbert, J. 2001b. Stochastic phonology. *GLOT International* 5(6). 195-207. Oxford: Blackwell publishers Ltd.

Pierrehumbert, J. 2002. Word-specific phonetics. *Laboratory Phonology VII.* 101-139. Berlin: Mouton de Gruyter.

Potts, C., Pater, J., Jesney, K., Bhatt, R. & Becker, M. 2010. Harmonic grammar with linear programming: from linear systems to linguistic typology. *Phonology* 27. 77-117.

Prince, A. & Smolensky, P. 1993. Optimality Theory: Constraint interaction in generative grammar. Ms., Rutgers University, New Bruswick, NJ., & University of Colorado, Boulder, CO.

Prince, A. & Tesar, B. 2004. Learning phonotactic distribution. In Kager, R., Pater, J., & Zonneveld, W. (eds.), *Constraints in Phonological Acquisition.* 245-291. Cambridge: Cambridge University Press.

Sadeniemi, M. 1949. *Metriikkamme Perusteet*. Helsinki: SKST.

Saffran, J. 2003. Statistical language learning: Mechanisms and constraints. *Current Directions in Psychological Science* 12.4. 110-114.

Saffran, J. & Thiessen, E. 2003. Pattern induction by infant language learners. *Developmental Psychology* 39.3. 484-494.

Shinohara, S. 1997. Default accentuation and foot structure in Japanese: Analysis of Japanese adaptation of French words. *MIT Working Papers in Linguistics* 30. 263-290.

Shinohara, S. 2000. Default accentuation and foot structure in Japanese: Evidence from Japanese adaptations of French words. *Journal of East Asian Languages* 9, 55-96.

Shinohara, S. 2004. Emergence of universal grammar in foreign word adaptation. In Kager, R., Pater, J., & Zonneveld, W. (eds.), *Constraints in Phonological Acquisition*. 292-320. Cambridge: Cambridge University Press.

Skousen, R. 1989. *Analogical Modeling of Language*. Dordrecht: Kluwer Academic Publishers.

Smolensky, P. 1996. The initial state and 'Richness of the Base' in Optimality Theory. Technical Report JHU-CogSci-96-4, Cognitive Science Department, Johns Hopkins University, Baltimore, MD. [ROA 86, http://roa.rutgers.edu]

Silva, D. 1992. *The Phonetics and Phonology of Stop Lenition in Korean.* PhD Dissertation, Cornell University, Ithaca: CLC Publications.

Silva, D. 2006. Acoustic evidence for the emergence of tonal contrast in contemporary Korean. *Phonology 23*. 287-308.

Tversky, A. 1977. Features of similarity. *Psychological Review 84*. 327-352.

Walker, R. 1995. Mongolian stress: Typological implications for nonfinality in unbounded systems. *Phonology at Santa Cruz* 4. 85-102.

Whalen, D.H. & Levitt, A. G. 1995. The universality of intrinsic F0 of vowels. *Journal of Phonetics 23*. 349-366.

Yang, B-G. 1996. A comparative study of American English and Korean vowels produced by male and female speakers. *Journal of Phonetics 24*. 245-261.

Zec, D. 1999. Footed tones and tonal feet: rhythmic constituency in a pitch-accent language. *Phonology* 16. 225-264.

Zuraw, K. 2000. *Patterned Exceptions in Phonology*. PhD dissertation, UCLA.