

## **UC Berkeley**

### **Proceedings of the Annual Meeting of the Berkeley Linguistics Society**

#### **Title**

Emergent Hidden Grammar: Stochastic Patterning in Korean Accentuation of Novel Words

#### **Permalink**

<https://escholarship.org/uc/item/5v6337c5>

#### **Journal**

Proceedings of the Annual Meeting of the Berkeley Linguistics Society, 36(36)

#### **ISSN**

2377-1666

#### **Author**

Kim, Hyun-Ju

#### **Publication Date**

2016

Peer reviewed

PROCEEDINGS OF THE THIRTY SIXTH ANNUAL MEETING OF THE  
**BERKELEY LINGUISTICS SOCIETY**

February 6-7, 2010

**General Session**

**Special Session**

Language Isolates and Orphans

**Parasession**

Writing Systems and Orthography

**Editors**

Nicholas Rolle  
Jeremy Steffman  
John Sylak-Glassman

Berkeley Linguistics Society  
Berkeley, CA, USA

Berkeley Linguistics Society  
University of California, Berkeley  
Department of Linguistics  
1203 Dwinelle Hall  
Berkeley, CA 94720-2650  
USA

All papers copyright © 2016 by the Berkeley Linguistics Society, Inc.

All rights reserved.

ISSN: 0363-2946

LCCN: 76-640143

# Contents

<b>Acknowledgments</b>	<b>v</b>
<b>Foreword</b>	<b>vii</b>
<i>Basque Genitive Case and Multiple Checking</i> Xabier Artiagoitia .....	1
<i>Language Isolates and Their History, or, What's Weird, Anyway?</i> Lyle Campbell .....	16
<i>Putting and Taking Events in Mandarin Chinese</i> Jidong Chen .....	32
<i>Orthography Shapes Semantic and Phonological Activation in Reading</i> Hui-Wen Cheng and Catherine L. Caldwell-Harris .....	46
<i>Writing in the World and Linguistics</i> Peter T. Daniels .....	61
<i>When is Orthography Not Just Orthography? The Case of the Novgorod Birchbark Letters</i> Andrew Dombrowski .....	91
<i>Gesture-to-Speech Mismatch in the Construction of Problem Solving Insight</i> J.T.E. Elms .....	101
<i>Semantically-Oriented Vowel Reduction in an Amazonian Language</i> Caleb Everett .....	116
<i>Universals in the Visual-Kinesthetic Modality: Politeness Marking Features in Japanese Sign Language (JSL)</i> Johnny George .....	129
<i>Equative and Predicational Copulas in Thai</i> Nancy Hedberg and David Potter .....	144

<i>On the Reflexive-Antipassive Polysemy: Typological Convergence from Unrelated Languages</i> Katarzyna Janic .....	158
<i>Position and Height Asymmetries in Hiatus Resolution: A Case Study of Korean VV Sequences</i> Hijo Kang .....	174
<i>Negative Concord in Western Armenian</i> Hrayr Khanjian .....	188
<i>Emergent Hidden Grammar: Stochastic Patterning in Korean Accentuation of Novel Words</i> Hyun-Ju Kim .....	203
<i>Evidentiality in Korean Conditional Constructions</i> Iksoo Kwon .....	218
<i>The Source-Goal Asymmetry in SLA</i> Wojciech Lewandowski .....	233
<i>Subject Relatives and Expletives in Early New High German</i> Caitlin Light .....	247
<i>An Embodied Account of Argument Structure Development</i> Josita Maouene, Nitya Sethuraman, Mounir Maouene, and Linda B. Smith .....	261
<i>A Gujarati Origin for Scripts of Sumatra, Sulawesi and the Philippines</i> Christopher Miller .....	276
<i>A Quantitative Analysis of Nominative/Genitive Alternation in Japanese</i> Satoshi Nambu .....	292
<i>A Sibling Precedence Approach to the Linearization of Multiple Dominance Structures</i> David Potter .....	307
<i>Surface Faithfulness Phenomena and the Consonantal Root in the Modern Hebrew Verb System</i> Tom Recht .....	322
<i>Partial Wh-Movement and Wh-Copying in Dutch: Evidence for an Indirect Dependency Approach</i> Ankelien Schippers .....	338
<i>Computational Cognitive Morphosemantics: Modeling Morphological Compositionality in Hebrew Verbs with Embodied Construction Grammar</i> Nathan Schneider .....	353

<i>Some Hypotheses About Possible Isolates within the Atlantic Branch of the Niger-Congo Phylum</i> Guillaume Segerer .....	368
<i>From Relativization to Nominalization and More: Evidence from the History of Okinawan</i> Reijirou Shibasaki .....	382
<i>A Cross-linguistic Study of Sound Symbolism: The Images of Size</i> Kazuko Shinohara and Shigeto Kawahara .....	396
<i>Testing for Frequency and Structural Effects in an English Stress Shift</i> Morgan Sonderegger .....	411
<i>Neighborhood Density in Phonological Alternations</i> Sverre Stausland Johnsen .....	426
<i>Person Indexicals in Uyghur Indexical Shifting</i> Yasutada Sudo .....	441
<i>Metathesis and Reanalysis in Ket</i> Edward Vajda .....	457
<i>An Empirical Investigation of Typicality and Uniqueness Effects on Article Choice in Attributive-Possession NPs</i> Gregory Ward, Christopher Ahern, and Tom Hayden .....	472
<i>Perception of Illegal Contrasts: Japanese Adaptations of Korean Coda Obstruents</i> James D. Y. Whang .....	488
<i>Diglossia versus Register: Discursive Classifications of Two Sinhala Varieties</i> Cala Zubair .....	499



# Acknowledgments

The editors of the 36th Annual Meeting of the Berkeley Linguistics Society are grateful to conference participants, our volunteers, session chairs, and the faculty, all of whom made the event an intellectually stimulating and enriching event. We would like to extend our sincere gratitude to the contributors of this volume for their professionalism, responsiveness, attention to detail, and patience in the editorial process, without which this would not have been possible.

Within our department, special thanks go to Paula Floro and Belén Flores for all their support with BLS 36 and this annual conference in general. The editors wish to also thank the executive committee of BLS 36 who organized and ran the conference, and Zachary O'Hagan for expertise in the final compilation using LaTeX.

Finally, we would like to thank the following organizations for their generous financial support:

Department of Linguistics  
Graduate Assembly  
Social Sciences Division  
Student Opportunity Fund  
International Computer Science Institute (ICSI)





# Foreword

This monograph contains 34 of the 51 talks given at the 36th Annual Meeting of the Berkeley Linguistics Society (BLS 36), held in Berkeley, California, February 6-7, 2010. The conference included a General Session, one Special Session entitled *Language Isolates and Orphans*, and one Parasession entitled *Writing Systems and Orthography*. It was planned and run by the second-year graduate students in the Department of Linguistics at the University of California, Berkeley. The members of this executive committee were Jessica Cleary-Kemp, Clara Cohen, Stephanie Farmer, Melinda Fricke, Laura Kassner, and John Sylak-Glassman.

The papers contained herein were edited principally for style by the three editors Nicholas Rolle, Jeremy Steffman, and John Sylak-Glassman, and then given back to contributors to make changes. Nicholas Rolle took upon primary editorial responsibilities, Jeremy Steffman was an undergraduate editorial assistant, and John Sylak-Glassman helped to edit papers. Upon the final resubmission, the final versions of these papers were incorporated by Zachary O'Hagan and Nicholas Rolle into the monograph found here. Our goal has been the speedy publication of these proceedings, and as such, certain aspects – e.g., the complete unification of formatting – have been sacrificed. It is our belief that this does not detract from the final publication in any way.

Nicholas Rolle  
Jeremy Steffman  
John Sylak-Glassman

January 2016



# Emergent Hidden Grammar: Stochastic Patterning in Korean Accentuation of Novel Words<sup>1</sup>

HYUN-JU KIM  
*Stony Brook University (SUNY)*

## 1 Introduction

The lexical accent pattern of a word in North Kyungsang Korean (NKK) is lexically determined, as illustrated in the following minimal triplet: *káci* ‘kind,’ *kací* ‘eggplant,’ *káci* ‘branch.’ Syllable structure does not categorically predict accent placement: for example, a heavy syllable does not necessarily attract accent (e.g. *camcári* ‘dragon,’ *Árísín* ‘older person,’ *susukkán* ‘kaoliang stalk’). However, it has been claimed that NKK speakers accent loanwords on the basis of syllable structure (Kenstowicz and Sohn 2001). How do NKK speakers decide to accent newly adopted words, which lack accent specification in UR? Do they accent new words randomly or in certain regular patterns? In order to answer these questions, an experimental study was performed using novel words to examine what determines accent placement when no lexical entry is available for accent.

Novel words have been used in many experimental studies recently. There have been at least two different views on how to explain patterns in novel words. One holds 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 phonetically similar to the novel words, or which are most frequent in the lexicon. Zuraw (2000) showed that Tagalog native speakers’ judgment of nasal substitution in novel words is consistent with the lexical frequency of existing words. Hayes and Londe (2006) also showed that Hungarian vowel harmony patterns in novel words are consistent with lexical frequency. 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

---

<sup>1</sup> I am indebted to Ellen Broselow and Marie Huffman for their valuable input and help with my study. All remaining errors are of my own. This paper was supported by NSF grant BCS-07460227 to Ellen Broselow, Marie Huffman, and Nancy Squires.

default stem-final stress is dominant (80%) in Russian stress placement in novel words, even though the occurrence of stem-final stress is much less (30%) in the lexicon.

This study presents empirical evidence that the accent patterns in novel words do not originate from analogy to phonetically similar familiar words. Rather, the accent pattern of novel words reflects statistical patterning in the lexicon. A corpus study showed that lexical distribution of NKK accent patterns is phonologically patterned: penultimate accent is common where all the syllables are light; final accent is more frequent where the final syllable is heavy. Lexical statistics revealed probabilistic structure-sensitive patterning in the lexicon even if exceptions obscure the patterning. This study will show that the accent patterns in novel words actually match this statistical patterning in the lexicon. This indicates that NKK speakers have implicit knowledge of the statistical patterning and apply it to novel words which lack lexical specification.

The rest of the paper is organized as follows: Section 2 presents a corpus study to examine statistical patterning in the lexicon. Section 3 presents experimental studies of NKK accentuation in novel words testing (i) the effect of analogy to phonetically similar words; (ii) stochastic patterning in novel words; and (iii) default accentuation. Section 4 discusses how accent patterns in novel words coincide with lexical statistical patterning. Section 5 is the conclusion.

## 2 Lexical Statistical Patterning

A corpus study was performed to examine lexical statistics based on a corpus of current Korean word usage (The National Institute of the Korean Language 2003; www.korean.go.kr). If the patterns which emerge in novel words are consistent with statistical patterning in the lexicon, this will support a stochastic grammar based on lexical frequency (Zuraw 2000; Hayes & Londe 2006).

The corpus consists of “pure” native words, excluding Sino-Korean words and loanwords. In order to control morphological effects on accentuation, only “monomorphemic” nouns were selected. Excluding loanwords, Sino-Korean words, derived nouns, and words longer than three syllables yielded 1275 bisyllabic and trisyllabic monomorphemic native Korean nouns out of 39,856 nouns. Of these 1275 words, 865 were bisyllabic and 410 were trisyllabic. Since the corpus does not indicate accent, three native speakers were asked to indicate the accent pattern on the existing words.<sup>2</sup>

As shown in (1), the bisyllabic words show a relatively balanced accentuation pattern (double 30% vs. penult 40% vs. final 30%).

---

<sup>2</sup> Native speakers disagreed on 8% of the words (58 of the 865 bisyllabic words and 45 of the 410 trisyllabic words, e.g. hwíp<sup>h</sup>áram ~ hwip<sup>h</sup>arám ‘whistle’; tórjállnim ~ torjállnim ‘brother-in-law’). For these words, the pattern which 2 out of 3 speakers assigned was chosen for the analysis.

*Stochastic Patterning in Korean Accentuation of Novel Words*

(1) Distribution of accent for bisyllabic words according to syllable weight combinations (The number given in parentheses is the percentage of each count)

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

\*H: heavy syllable; L: light syllable

However, accent patterns are not distributed evenly across syllable weight combinations. The distribution shows more frequent penultimate accent in LL (49%), more frequent double accent in HL (42%) compared to 26% in LL and 23% in LH, more frequent final accent in LH (40%) compared to 25% in LL and 18% in HL, and less common penultimate accent in HH (28%) than in other combinations. Statistical analyses using Pearson's chi-square test confirmed that the distribution of accent was significantly different according to syllable weight combinations (LL vs. HL:  $\chi^2(2)=12.16$ ,  $p=.002$ ; LL vs. LH:  $\chi^2(2)=13.67$ ,  $p=.001$ ; LL vs. HH:  $\chi^2(2)=15.84$ ,  $p<.001$ ; HL vs. HH:  $\chi^2(2)=13.4$ ,  $p=.001$ ). In sum, the accent distribution in bisyllabic words showed a clear tendency toward accented heavy syllables.

Unlike the overall even distribution in bisyllabic words, lexical statistics of trisyllabic words showed uneven accent distribution: 306 out of 411 words (75%) received penultimate accent, 66 words were assigned double accent (16%), 29 words were given final accent (7%), and only 10 words received antepenultimate accent (2%) overall in trisyllabic words as shown in (2).

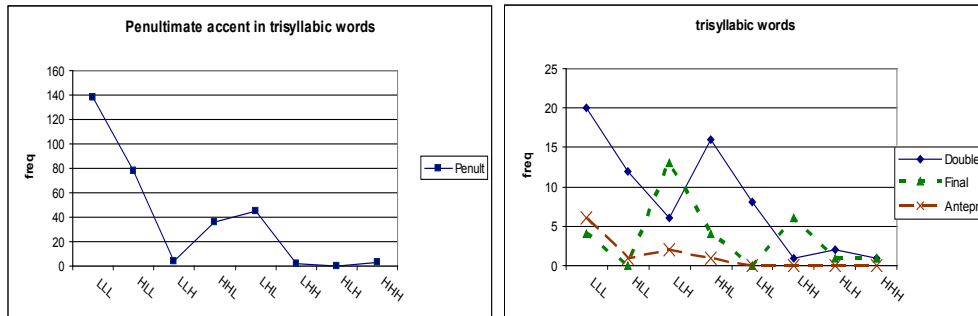
(2) Distribution of accent for trisyllabic words according to syllable weight combinations

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

The results showed that the preference for penultimate accent was overwhelmingly high (75%) in trisyllabic words, while the number of trisyllabic words with other accent patterns was fairly small: double accent 16%; final 7%; antepenultimate 2%. Despite the predominance of penultimate accent, Pearson's chi-square test showed a significant association between accent type and syllable weight combination. Final accent was the most preferred in words ending in a word-final heavy syllable (52% in LLH; 67% in LHH), which was significantly different from LLL combination ( $\chi^2(3)=77.16$ ,  $p<.001$ ), whereas final accent was one of the least preferred accents in trisyllabic words (7%). This demonstrates that final accent is chosen not accidentally but because accent favors the word-final heavy syllable. In addition, double accent in HHL was significantly more frequent than in LLL ( $\chi^2(3)=11.96$ ,  $p=.008$ ). However, a word-initial heavy syllable (HLL) itself did not affect accent placement because the pattern of HLL was not significantly different from that of LLL combinations ( $\chi^2(3)=3.67$ ,  $p=.3$ ). Double accent in HLL (13%:12 out of 91 HLL words) was significantly less frequent than double accented words (28%) in HHL ( $\chi^2(3)=12.92$ ,  $p=.005$ ).

The following graphs in (3) were based on the raw frequency of the results. The left-hand graph illustrates the occurrence of penultimate accent in LLH, LHH, HLH, and HHH.

(3) Structure-sensitive accent distribution in trisyllabic words



While penultimate accent was prevalent in LLL and HLL words, it was much rarer in words ending in heavy syllables. As illustrated in the right-hand graph in (3), double accent was more frequent in HHL and final accent in LLH and LHH.

To summarize, the lexical statistics showed that penultimate accent is the most frequent in the lexicon. However, the distribution of accent according to syllable weight combination revealed structure-sensitive patterning: there is a strong tendency toward final accent on LH and LLH; a tendency toward double accent on HH and HHL. The next section addresses the question of whether NKK speakers actually internalize the statistical patterning and apply it to novel words.

### **3 Experiments**

An experimental study of NKK accentuation was performed using novel words to examine the factors determining accent placement when no lexical entry is available with an accent specification. Three hypotheses guided this study. (1) The Analogy Hypothesis: accent patterns in novel words will follow the patterns of existing words that are phonetically similar to the novel words. Analogy effects have been found in Spanish stress placement with novel words (Eddington 2000, 2004; Face 2004). Face (2004) especially showed an effect of analogy to phonetically similar words, and effects of lexically subregular patterns and morphological category in nonsense words. Therefore, in the absence of lexical accent specification, it is possible that speakers access neighboring lexical items that are similar to the novel words and follow the patterns of these similar words. (2) The Stochastic Accent Hypothesis: accentuation in novel words will be consistent with statistical patterning in the lexicon. (3) 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, Kim (1997) actually proposes default penultimate accent for underlyingly accentless words. Kim's hypothesis predicts penultimate accent as default regardless of syllable structure since native accentuation cannot be solely determined by syllable structure.

#### **3.1 Experiment 1: Analogy Test**

In this experiment, the Analogy Hypothesis is tested using nonce words similar to existing words. This hypothesis predicts that patterns in nonce words will follow those of words that are phonetically similar to the target word.

##### **3.1.1 Participants**

12 subjects, ranging in age from 28 to 55 years (mean 41), participated in this study. Participants were born and grew up in Daegu, a city located in the northern part of Kyungsang province, and most of them still lived in the city at the time of the experiments. Five were male and seven were female. They were paid for their participation, which took less than half an hour.

##### **3.1.2 Materials**

The stimulus set consisted of nonce words, 24 bisyllabic and 26 trisyllabic. The nonce words were identical to existing words except for one consonant in the final syllable, of which difference is one place feature from the corresponding consonant in a real word; for example, *tonsem*, *hamjipat* were constructed by changing a place feature of the final consonant of real words, *tonséŋ* 'younger sibling,'




*hamjipák* ‘wooden bowl.’ Forms ending with a vowel were created by changing onset consonants of the final syllable (e.g. *kampa*, *mujide* from *kamjá* ‘potato,’ *mújike* ‘rainbow’). Following the finding that the onset of a word activates a set of lexical candidates (Allopenna et al. 1998; Marslen-Wilson 1987), I assumed the nonce stimuli were similar enough to activate neighboring lexical items. An equal number of nonce forms was given for each accent pattern and for each combination of light (open) and heavy (closed) syllables.<sup>3</sup>

### 3.1.3 Procedure

Participants were recorded in a quiet room 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, for example “*atiri* \_\_\_\_\_ *to satalla k<sup>h</sup>atara*” (My son asked me to buy \_\_\_\_\_ too) with 2 repetitions. The second readings were analyzed. Nonce words were presented with pictures of unfamiliar cartoon characters as in (4) and introduced as the names of new cartoon characters which had recently become popular among children.

(4) Example sentence

atiri  *kampa* to satalla k<sup>h</sup>atara  
 ‘My son asked me to buy *kampa* 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, accent decision was made based on the pitch contour generated using Praat.

### 3.1.4 Results

A total of 599 tokens (50 items x 12 subjects, excluding one token with production error) were collected for the analogy test. The distribution of actual accent patterns was compared to that of analogy accent patterns using a loglinear analysis. The independent variables were syllable weight combination, e.g. LL, HL, LH, HH (H:heavy; L:light), accent type (double; penultimate; final)<sup>4</sup>, and accent

<sup>3</sup> The list of all test words used in the study appears in the Appendix.

<sup>4</sup> Initial accent was rarely produced in novel words. Only two tokens were produced in initial accent (one in analogy set and the other in non-analogy set). Therefore, it was omitted in statistical

*Stochastic Patterning in Korean Accentuation of Novel Words*

word group (actual vs. analogy). The dependent variable was the count of tokens for each variable.

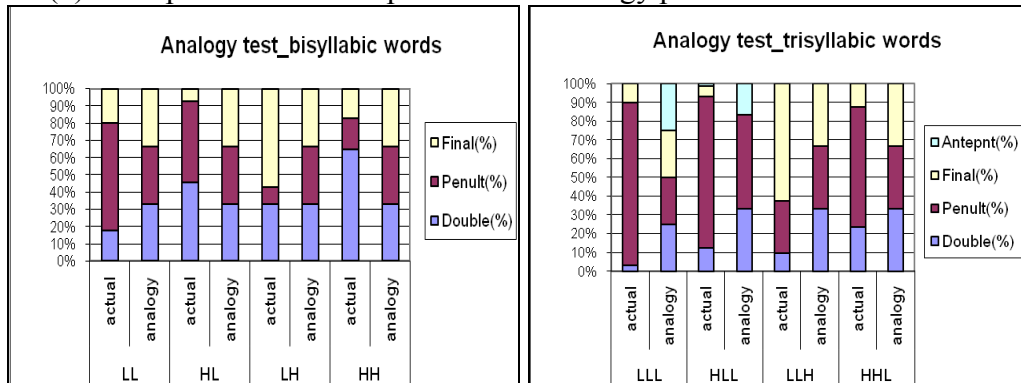
The results showed that only 41% of responses (246 out of 598 tokens) showed analogy patterns, while 59% showed no effect of analogy. As shown in (5), penultimate accent was the most frequent overall (51.2%; 306 out of 598 tokens). Penultimate accent was dominant in words with only light syllables (76.2%) and in forms with an initial heavy syllable (64.3%). The predominance of penultimate accent did not comply with the result predicted by the Analogy Hypothesis, which predicted that only 33% of these forms should have penultimate accent. The distribution of final accent was also different from that predicted by analogy: 86 out of 144 tokens (59.7%) were final accented, as opposed to 33% predicted by analogy. For the words with two initial closed syllables (HH/HHL), double accent was more common than in words with other syllable weight combinations.

(5) Overall results of actual accent patterns

	Double (%)	Penult (%)	Final (%)	Total (%)
LL/LLL	16 (9.5)	<b>128 (76.2)</b>	24 (14.3)	168 (100)
HL/HLL	42 (29.4)	<b>92 (64.3)</b>	9 (6.3)	143 (100)
LH/LLH	31 (21.5)	27 (18.8)	<b>86 (59.7)</b>	144 (100)
HH/HHL	<b>63 (44.1)</b>	59 (41.3)	21 (14.7)	143 (100)
Total	152 (25.4)	<b>306 (51.2)</b>	140 (23.4)	598 (100)

The distribution of actual accent patterns compared to that of analogy patterns is illustrated in the following figures.

(6) Comparison of actual patterns and analogy patterns



If analogy influences accent patterns in nonce words, actual patterns with double, penult, and final accent are expected to be evenly distributed in each syllable

analyses.

weight combination, as in the analogous real words. However, the results showed that actual patterns did not follow analogy patterns. For instance, more final accent in LH and more double accent in HH were found in actual patterns than in analogy patterns. A loglinear analysis confirmed that the distribution of actual patterns was significantly different from that of analogy patterns ( $\chi^2(3) = 84.28$ ,  $p < .001$ ). It also showed that interaction between syllable weight combination and accent type was significant ( $\chi^2(9) = 427.72$ ,  $p < .001$ ).

To summarize, the analogy test showed that the actual patterns were significantly different from analogy patterns to similar words. This result suggests that analogy to phonetically similar words is irrelevant to NKK accentuation of novel words.

### 3.2 Experiment 2: Stochastic Accent and Default Accent Test

Experiment 2 was designed to test whether the accent patterns of novel words reflect the statistical patterning in the lexicon (Stochastic Accent Hypothesis) or default penultimate accentuation (Default Accent Hypothesis). The Stochastic Accent Hypothesis predicts that heavy syllables will tend to attract accent in novel words, in accordance with the statistical association between accent and syllable weight in the lexicon. On the other hand, the Default Accent Hypothesis predicts that penultimate accent will be the most common regardless of syllable structure.

#### 3.2.1. Participants and Procedure

The participants were the same as those in Experiment 1 (12 NKK speakers, ranging in age from 28 to 55 years (mean 41), five male and seven female). Experiment 2 was run following the same procedure as in Experiment 1.

#### 3.2.2 Materials

The test word set consisted of four possible combinations of syllables for bisyllabic words and for trisyllabic words: LL, HL, LH, HH, 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 within a word; for example, only the vowel *a* was used as in a word *takapa* or the vowel *i* was used as in a word *citiki*. 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* and *n* (e.g. *mana*). 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. *tapnaka* → *tamnaka*). Four

or five items were included for each syllable weight combination. In total, 36 words were created: 20 bisyllabic words and 16 trisyllabic words.<sup>5</sup>

### 3.2.3 Results

A total of 427 tokens were collected from 36 test words with 12 subjects excluding 5 tokens with production errors<sup>6</sup>. 237 tokens were bisyllabic and 190 tokens were trisyllabic. The Default Accent Hypothesis predicts that penultimate accent would be the most frequent regardless of syllable structure. As shown in (7), overall results showed that penultimate accent was the most frequent in bisyllabic and trisyllabic words: 45% in bisyllabic words; 56% in trisyllabic words. However, double accent was assigned frequently in bisyllabic words (37%) although it was not as common as penultimate accent. Final accent was least favored (less than 20%) in both groups (18% in bisyllabic words; 16% in trisyllabic).

(7) Overall results in bisyllabic words and in trisyllabic words

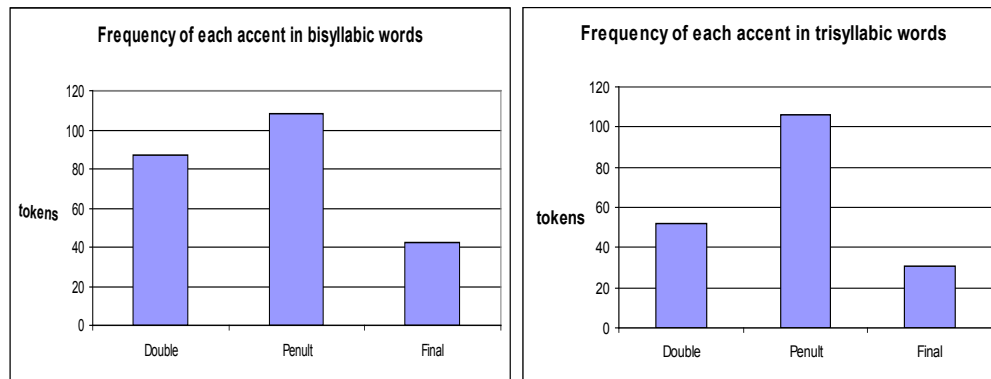


Table (8) presents counts of each accent according to syllable weight combinations in bisyllabic words.

(8) Distribution of accent for bisyllabic words according to syllable weight combinations

	Double (%)	Penult (%)	Final (%)	Total
LL	10 (16.7)	<b>49 (81.7)</b>	1 (1.7)	60
HL	15 (25.4)	<b>42 (71.2)</b>	2 (3.4)	59
LH	17 (28.8)	11 (18.6)	<b>31 (52.5)</b>	59
HH	<b>45 (76.3)</b>	6 (10.2)	8 (13.6)	59
Total	87 (36.7)	<b>108 (45.6)</b>	42 (17.7)	237

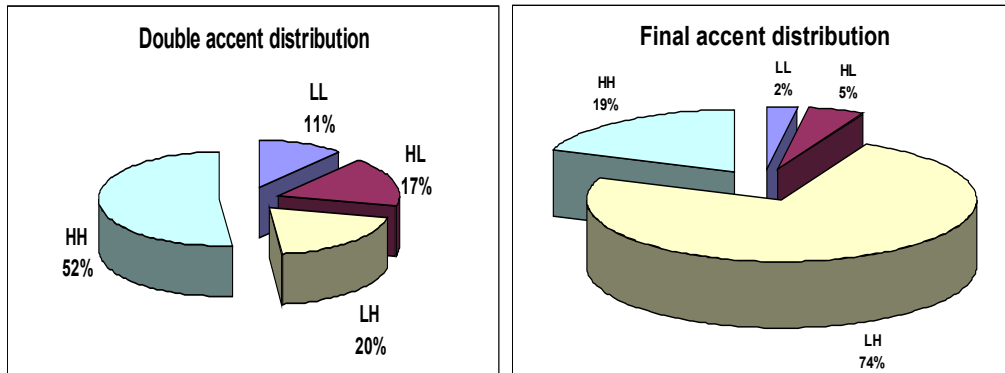
<sup>5</sup> The list of all test words used in the study appears in the Appendix.

<sup>6</sup> The tokens excluded due to production errors were as follows: *kakpap* produced for the target word *kakpa* (S2); *manaj* for *manhaj* (S2); *tampa* for *tapam* (S7); *kintinci* for *kikdinci* (S7); *tanjkanpa* for *tapkapa* (S11).

Penultimate accent was chosen more than 80% of the time in words with all light syllables (LL). When the initial syllable was heavy (HL), accent placement was similar to that of words with all light syllables (LL): penultimate accent for HL was still as high as 71%. However, the weight of the final syllable was crucial in inducing non-penultimate accent. The final syllable was accented in 53% of the words with a final heavy syllable, which was significantly different from patterns of the other combinations: final accent was produced in fewer than 5% of LL and HL combinations ( $p < .001$ )<sup>7</sup>. Double accent assignment increased to 76% in words beginning with heavy syllables when the final syllable was heavy. This result contrasted with the low assignment of double accent in HL combination (25%). Therefore, the final heavy syllable was a trigger for non-penultimate accent placement.

Figure (9) presents more clearly the distribution of double accent and final accent according to syllable weight combinations, which demonstrates structure-sensitive accentuation.

(9) Structure-sensitive accentuation in bisyllabic words



Among the words with double accent, the HH combination was the most favored and the percentage (52%) was significantly higher than that of other combinations (LL, HL, LH;  $p < .05$ ). In addition, most of the final accented words were LH. The probability of final accent in LH (74%) was significantly higher than in LL (2%) ( $p < .001$ ). Even though final accent was not favored, final accent was chosen when the word-final syllable was heavy.

As 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. Double accent was also not preferred in most combinations except in words where the initial two syllables were heavy (HHL). The results are shown in (10).

<sup>7</sup> The results were compared using a repeated measures analysis of variance (ANOVA). The independent variables used were syllable weight combination and accent position. The dependent variable was the number of tokens.

*Stochastic Patterning in Korean Accentuation of Novel Words*

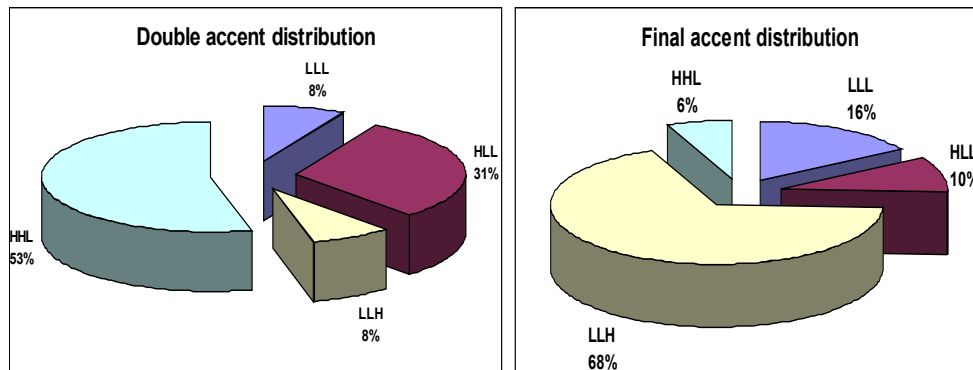
(10) Distribution of accent for trisyllabic words according to syllable weight combinations

	Double (%)	Penult (%)	Final (%)	Total
LLL	4 (8.3)	<b>39 (81.3)</b>	5 (10.4)	48
HLL	16 (34)	<b>28 (59.6)</b>	3 (6.4)	47
LLH	4 (8.3)	<b>22 (45.8)</b>	<b>21 (43.8)</b>	48
HHL	<b>28 (59.6)</b>	17 (36.2)	2 (4.3)	47
Total	52 (27.4)	<b>106 (55.8)</b>	31 (16.3)	190

Penultimate accent was dominant in the combinations LLL (81%) and HLL (60%). However, penultimate accent was less frequent in the combinations LLH and HHL: 46% and 36% respectively. Instead, final accent became more common in LLH (44%) and double accent in HHL (60%) compared to LLL and HLL combinations. The tendency toward final accent for LLH (44%) was significantly stronger than the tendency toward final accent in other combinations (10% in LLL ( $p=.019$ ); 6% in HLL ( $p=.02$ ); 4% in HHL ( $p=.002$ )). Even though double accent was more common in HLL (34%) than in LLL (8%), the difference did not reach statistical significance. However, double accent increased to 60% in HHL, which was significantly higher than the number of double accented words in all the other combinations ( $p<.05$ ).

The structure-sensitive accentuation was revealed more clearly in the distribution of each accent according to syllable weight combinations as illustrated in (11).

(11) Structure-sensitive accentuation in trisyllabic words



Words with double accent were significantly more common in HHL than in other combinations ( $p<.05$ ). Also, 68% of the words with final accent were LLH combinations, which was significantly higher than those in other combinations ( $p<.05$ ).

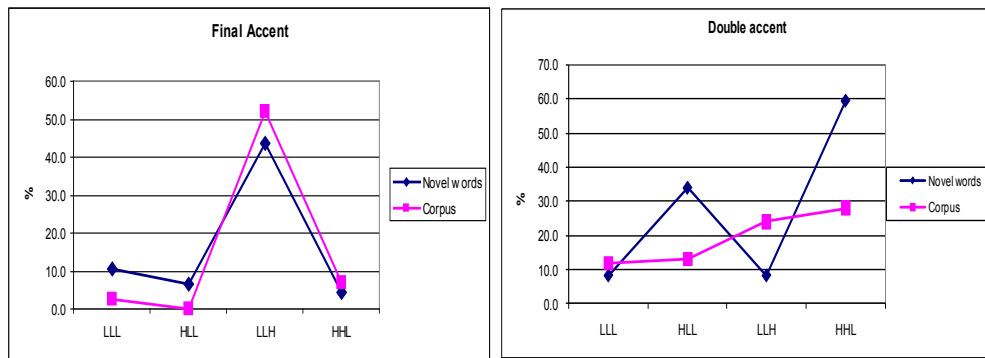
To summarize, penultimate accent was the most frequent in nonce forms. However, accentuation in nonce forms was also syllable structure-sensitive, contrary to the prediction of the Default Accent Hypothesis. The structure-sensitive

patterns manifested in nonce words actually complied with statistical patterning in the lexicon, which will be discussed in the following section.

#### 4 Discussion

The results showed that analogy was not a factor in accentuation of nonce words. The accent regularity expressed in nonce words suggests that phonologically-grounded sources, not random choices, determine the patterns: both penultimate accent and structure-sensitive accent patterns. Penultimate accent was the most frequent; however, default accentuation was not sufficient to account for nonce word accentuation because it failed to predict structure-sensitive accent in nonce words. The source of the structure-sensitive patterns can be ascribed to statistical patterning in the lexicon. As Figure (12) shows, the accent patterns in the corpus and in nonce words match fairly well.

(12) Comparison of accent patterns in the corpus and in nonce words



Final accent was quite rare in LLL, HLL, and HHL but common in LLH, both in nonce words and in the corpus, as shown in the left-hand graph. The strong tendency toward accented word-final heavy syllable in nonce words accords with lexical frequency. The right-hand graph illustrates that double accent was preferred in HHL both in nonce words and in the corpus. The preference for double accent for HHL was even stronger in nonce words. On the other hand, double accent in HLL was not as common as in HHL in the corpus whereas double accent emerged rather strongly in HLL nonce words.

The structure-sensitive accentuation in nonce words is consistent with the statistical patterns in the lexicon. Furthermore, the structure-sensitive patterning appeared even more strongly in nonce words. This suggests that NKK speakers have implicit knowledge of the structure-sensitive patterning, even though it is obscured by exceptions. Weight-sensitive patterns of loanwords also provide another source of evidence that NKK speakers apply the knowledge of structure-sensitive patterning to newly adopted words: penultimate accent is generally assigned in words without heavy syllables (e.g.  $\text{ʃik}^h\text{áko}$  ‘Chicago’), final accent in words end-

ing in a heavy syllable (e.g. *si<sup>h</sup>ak<sup>h</sup>iŋ* 'stocking'), and double accent is preferred in words beginning with heavy syllables (e.g. *si<sup>h</sup>mp<sup>h</sup>óni* 'symphony'). 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.<sup>8</sup> However, these results suggest that the accent patterns in loanwords are in accordance with the native language association between syllable weight and accent.

## 5 Conclusion

An experimental study using novel words revealed a strong tendency toward accented heavy syllables as well as a preference for penultimate accent. A corpus study showed that these patterns are present in statistical patterning in the lexicon. This suggests that the syllable structure sensitivity and the predominant penultimate accentuation emerging in novel words are actually the reflection of statistical patterning in the lexicon. Furthermore, the results of an analogy test showed that analogy is not a factor in accentuation of novel words. I conclude that NKK speakers internalize lexical statistical patterning and that a stochastic grammar based on lexical frequency guides the accentuation of newly adopted words in NKK (Zuraw 2000; Hayes & Londe 2006).

## Appendix

### Experiment 1

CV.CV	CVC.CV	CV.CVC	CVC.CVC
suga	nikbe	gasiŋ	ʌkc <sup>h</sup> ʌp
cagu	damge	badap	bʌpsʌt
cop <sup>h</sup> a	cipde	giriŋ	hʌtgaŋ
gubu	gukda	hobat	ʌlgun
geni	nemgi	ginim	dangoŋ
nadi	gamba	gusʌt	bancup
CV.CV.CV	CVC.CV.CV	CV.CV.CVC	CVC.CVC.CV
acire, gamari	nalc <sup>h</sup> isi	jadinim	baŋraŋi
sat <sup>h</sup> umi, mucide	simmami	ʌriŋ	gombange
mudʌbi	doŋc <sup>h</sup> iri	arimaŋ	caŋttaŋgi
sadami	dalgugi	susuttaŋ	incʌlri
doragi	danmugi	gʌdirinŋ	cindamme
nʌsime	coŋami	ttiŋgegil	mindinne

<sup>8</sup> Accent placement in loanwords is not necessarily consistent with stress placement in the source language (e.g. *si<sup>h</sup>ak<sup>h</sup>iŋ* vs. *stócking*).



**Experiment 2**

CV.CV	CVC <sub>s/o</sub> .CV	CV.CVC <sub>s/o</sub>	CVC <sub>s/o</sub> .CVC <sub>s/o</sub>
mana	maŋna	manaŋ	maŋnaŋ
kapa	kakpa	kapak	kakpap
tapa	tampa	tapam	tampaŋ
cipi	cimpi	cipim	cimpin
tiki	tipki	tikik	tipkik

CV.CV.CV	CVC <sub>s/o</sub> .CV.CV	CV.CV.CVC <sub>s/o</sub>	CVC <sub>s/o</sub> .CVC <sub>s</sub> .CV
takapa	tapkapa	takapak	tapkanpa
pakaca	paŋkaca	pakacaŋ	paŋkanca
kitici	kiktici	kiticik	kiktinci
pipici	pimpici	pipicim	pimpinci

**References**

- Albright, A. & Hayes, B. 2001. Rules vs. Analogy in English Past Tenses: A Computational/Experimental Study, Ms. UCLA.
- Alloppenna, 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.
- Crosswhite, K., Alderete, J., Beasley, T., & Markman, V. 2003. Morphological Effects on Default Stress in Novel Russian Words. In G. Garding and M. Tsujimura (eds.), *WCCFL 22 Proceedings*, 151-164. Somerville, MA: Cascadilla Press.
- Eddington, D. 2000. Spanish Stress Assignment within the Analogical Modeling of Language. *Language* 76: 92-109.
- Face, T. 2004. Perceiving What Isn't There: Non-Acoustic Cues for Perceiving Spanish Stress. In Timothy L. Face (eds.), *Laboratory Approaches to Spanish Phonology*, 117-141. Mouton de Gruyter, Berlin.
- Hayes, B. & Londe, Z. 2006. Stochastic Phonological Knowledge: The Case of Hungarian Vowel Harmony. *Phonology* 23, 59-104. Cambridge University Press.

*Stochastic Patterning in Korean Accentuation of Novel Words*

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

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

Marslen-Wilson, W. 1987. Functional Parallelism in Spoken Word Recognition. *Cognition* 25, 71-102.

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

Hyun-ju Kim  
Stony Brook University (SUNY)  
Department of Linguistics  
S205 Social and Behavioral Sciences  
Stony Brook, NY 11794-4376

[hjkim@sunykorea.ac.kr](mailto:hjkim@sunykorea.ac.kr)