1. Introduction
The typical data patterns employed in traditional analysis within generative phonology are systematic with no variation. Even when exceptional or variant output forms may be attested, they have often been ignored in formal analysis. However, the variable, gradient, patterns recently attract much attention with the understanding that the exceptions and variants occur in a systematic way in many cases. The results of some recent research on the variation data (Bybee 2001; Boersma and Hayes 2001; Zuraw 2000, 2002, 2005; Albright 2002 a,b, 2005a, b; Albright and Hayes 2003; Pierrehumbert 2003; Hayes and Londe 2006) show that the variation pattern observed in the speakers’ behavior is matched by the statistical pattern in the lexicon, supporting the hypothesis that the speakers may internalize the variable lexical pattern and use the knowledge in their behavior. The important evidence in favor of the hypothesis comes from the data of productivity-testing such as loanword adaptation and Wug-test (Burko 1958).

The present study provides additional supporting evidence for the hypothesis, discussing the Korean case. It has been observed in the literature on Korean phonology and morphology that coronal stem-final obstruents of nouns are in free variation with [s] in prevocalic position (Ko 1989; Martin 1992; Hayes 1998 among others): for example, /patʰ-il/ [patʰil] ~ [pasil] ‘field, accusative’. Recent survey studies (Kim 2003; Choe 2004; and Kang et al. 2004) reveal that a wide set of obstruents [s, čʰ, ᵇʰ, č, t] may occur as free variants. Moreover, there is an order of usage preference among them: in general, s >> čʰ >> ᵇʰ >> č, t.

This paper first shows that the observed relative preference among variants is matched by the distribution of final obstruents of the noun stems (cf. Albright 2005a). This frequency-matching is not confined to nouns. Employing a certain set of Korean bound stems as experimental stimuli, I have conducted a judgment survey, which may be considered as a type of Wug-test. The survey results show the same type of stem-final obstruent variation as observed with the noun stems. The relative acceptability among variants is also matched by the distribution of noun-final obstruents. The observed frequency-matching suggests that native Korean speakers are aware of lexical frequency, and they use this lexical knowledge in their judgment task. Thus, this paper provides evidence in favor of the hypothesis about the internalization of the variable lexical pattern.

Further, unlike in most previous analyses, I argue that a correct analysis of Korean stem-final variation should employ an output-oriented generalization. This argument is based on the finding that the same type of variation is attested in two distinct
morphological categories, i.e., nouns and bound stems, with different paradigm systems. Most previous analyses of Korean noun stem-final variation (Ko 1989; Kang 2003a,b; Albright 2005a) are source-oriented, as it is specified how to derive a prevocalic variant form from its corresponding isolation form. But, this type of source-oriented analysis cannot be extended to the variation of the bound stem mainly because the isolation form of the bound stem does not exist.

The remainder of this paper is organized as follows. In the next section, I first discuss the general phonological processes of Korean that cause basic stem-final alternations in the noun paradigm. This is followed by a discussion of the recent survey studies which investigate patterns of variation involved in the alternations. In section 2, I show that all the observed patterns are consistent with frequency facts of Korean nouns. In section 3, I describe a judgment survey whose results show that topicalized bound stems exhibit similar stem-final obstruent variation. In section 4, I discuss how to analyze the findings of the present study, providing a possible formal analysis couched within the stochastic extension of Optimality Theory (Boersma and Hayes 2001; Zuraw 2005; Hayes and Londe 2006). In section 5, I show that all alternative analyses (Ko 1989; Kang 2003a,b; Albright 2005a,b; Park 2006) are subject to serious problems. In the final section, I will summarize conclusions of the present study.

1.1 Stem-final alternations in nouns

In Korean, obstruents show a three-way laryngeal distinction, lenis, aspirated and tense (or glottalized), as is shown below.

(1) Three-way laryngeal distinction among obstruents in Korean

<table>
<thead>
<tr>
<th></th>
<th>Labial</th>
<th>Coronal</th>
<th>Velar</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Stop</td>
<td>Fricative</td>
<td>Affricate</td>
</tr>
<tr>
<td>Lenis</td>
<td>p</td>
<td>s</td>
<td>č</td>
</tr>
<tr>
<td>Aspirated</td>
<td>pʰ</td>
<td>tʰ</td>
<td>čʰ</td>
</tr>
<tr>
<td>Tense</td>
<td>p’</td>
<td>t’</td>
<td>s’</td>
</tr>
</tbody>
</table>

These obstruents may be targeted in some phonological changes, two of which are crucial to the purpose of this paper. First, aspirated and tense obstruents in coda position become their homorganic lenis stop counterparts (marked in boldface in the above table). This coda neutralization applies with no exception. In addition, coronal obstruents are palatalized before the high front vowel [i]. Specifically, alveolar stops [t, tʰ] become palato-alveolar affricates [č, čʰ], and alveolar fricatives [s, s’] become palato-alveolar counterparts [ʃ, ʃ’] in the palatalization context. The former palatalization process applies only when a morpheme boundary intervenes whereas the latter is an automatic process.

These phonological processes cause alternations in the noun paradigm. As shown in (2), Korean noun stems may appear with a variety of final obstruents when they combine with certain vowel-initial suffixes such as the locative case marker –e.
However, isolation forms may end only with (unreleased) lenis stops, due to the coda neutralization. Also, before the nominative case marker -i, stem-final coronal obstruents are realized only as palato-alveolars such as [č, čʰ, ţ], due to the coronal palatalization. As a result, noun stems ending with obstruents show stem-final alternations within the paradigm. For instance, the final consonant of the noun stem meaning ‘field’ appears as [tʰ] before the locative case marker -e, [t] in the isolation form, and [čʰ] before the nominative case marker –i.

(2) Alternations of stem-final obstruents in nouns.

<table>
<thead>
<tr>
<th>stem-final C</th>
<th>(i) locative (-e)</th>
<th>(ii) isolation form</th>
<th>(iii) nominative (-i)</th>
<th>gloss</th>
</tr>
</thead>
<tbody>
<tr>
<td>Coronal</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>os-e</td>
<td>ot’</td>
<td>of-i</td>
<td>‘clothes’</td>
</tr>
<tr>
<td></td>
<td>patʰ-e</td>
<td>pat’</td>
<td>pacʰ-i</td>
<td>‘field’</td>
</tr>
<tr>
<td></td>
<td>nač-e</td>
<td>nat’</td>
<td>nač-i</td>
<td>‘day’</td>
</tr>
<tr>
<td></td>
<td>pieʰ-e</td>
<td>pit’</td>
<td>pieʰ-i</td>
<td>‘light’</td>
</tr>
<tr>
<td>Labial</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>pap-e</td>
<td>pap’</td>
<td>pap-i</td>
<td>‘rice’</td>
</tr>
<tr>
<td></td>
<td>ipʰ-e</td>
<td>ip’</td>
<td>ipʰ-i</td>
<td>‘leaf’</td>
</tr>
<tr>
<td>Dorsal</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>kuk-e</td>
<td>kuk’</td>
<td>kuk-i</td>
<td>‘soup’</td>
</tr>
<tr>
<td></td>
<td>pakʰ-e</td>
<td>pak’</td>
<td>pakʰ-i</td>
<td>‘outside’</td>
</tr>
<tr>
<td></td>
<td>puakʰ-e</td>
<td>puak’</td>
<td>puakʰ-i</td>
<td>‘kitchen’</td>
</tr>
</tbody>
</table>

In standard rule-based analysis of these stem-final alternations, the output form occurring before the vowel-initial suffixes (except for [i]-initial ones) is posited as the underlying form, (which is reflected in standard Korean orthography). The above mentioned phonological rules apply to the underlying form to produce unreleased stops and palatalized consonants in the isolation and nominative forms respectively. This standard rule-based analysis can be illustrated by the following derivation of some allomorphic forms of the stem /patʰ/.

(3) Standard rule-based analysis of the alternations.

<table>
<thead>
<tr>
<th>locative</th>
<th>isolation form</th>
<th>nominative</th>
</tr>
</thead>
<tbody>
<tr>
<td>Underlying form: /patʰ+e/</td>
<td>/patʰ/</td>
<td>/patʰ+i/</td>
</tr>
<tr>
<td>Coda neutralization</td>
<td>t</td>
<td></td>
</tr>
<tr>
<td>Palatalization</td>
<td>čʰ</td>
<td></td>
</tr>
<tr>
<td>Surface form: patʰe</td>
<td>pat</td>
<td>pacʰi</td>
</tr>
</tbody>
</table>
1.2 Variation in nouns

It has been observed in the literature (Ko 1989; Martin 1992; Hayes 1998; Y. Kang 2003a,b; Albright 2005a,b; and others) that pre-voc allic allomorphs of the noun stems, such as [patʰ-e] and [pačʰ-i] in (3), show free variation of final obstruents. The variants differ depending on the place of articulation of the obstruents. Coronal obstruents are mostly in free variation with [s], less frequently with [cʰ]: for example, /patʰ-il/ [patʰil] ~ [pasɪl] ~ [pačʰ-il] ‘field, accusative’. Frequent occurrence of the final variant [s] has been extended to the pronunciation of the inflected forms of English loanwords ending with coronal stops. Word-final /t, d/ in English are normally realized as [s] in Korean when they are combined with vowel-initial suffixes: for instance, [kus-in] ‘good, topicalized’ (Ko 1989). Noncoronal aspirated/fortis stops, [pʰ, kʰ, k’], are in free variation with their homorganic lenis counterparts, [p, k]: for example, /ipʰ-e/ [ipe] ‘leaf, locative’ and /pskʰ-e/ [pske] ‘kitchen, locative’.

Some recent studies (Kim 2003; Choe 2004; and Kang et al. 2004) reveal that the variation patterns involved are much more complicated than previously thought. In all these studies, a sizeable pool of native Seoul Korean speakers were employed: 1,174 subjects in Choe, 156 subjects in Kang et al., and 350 subjects in Kim. Subjects were consulted about their actual and preferred pronunciations in the surveys by Choe and E. Kang et al. respectively. Kim’s study, which reports somewhat limited data on noun stem-final variations, is based on the phonetic transcriptions of the speech of subjects who read experimental sentences.

Most notably, the range of the coronal variants is quite wide, as can be seen in (4). Five coronal obstruents [s, cʰ, tʰ, c, t] are involved as free variants. For instance, in (4bv), the locative forms of the stem /tוסʰ/ may have all the five coronal obstruents [s, cʰ, tʰ, c, t] in the stem-final position. Notice that most other noun stems have variants besides the one ending in [s] and the “standard” phonetic output of the underlying form which is either identical to the underlying form or, before [i]-initial suffixes, the result of the cross-morphemic coronal palatalization.

(4) Noun variations: Stem-final obstruent = coronal (data selected from Kim (2003), Choe (2004), and E. Kang et al. (2004))

| Stem-final C | UR       | Standard | Variants                  | gloss     | Source
|--------------|----------|----------|---------------------------|-----------|-------
| a. /tʰ/      | (i) /patʰ-i/ | pacʰ-i   | pas-i [paʃi]               | ‘field’   | Choe  |
|              | (ii) /patʰ-il/ | patʰ-il  | pas-il ~ pacʰ-il          | ‘field’   | Choe  |
|              | (iii) /soṭʰ-e/ | soṭʰ-e   | sos-e ~ sočʰ-e ~ so-t-e   | ‘pot’     | Kang  |
|              | (iv) /miṭʰ-il/ | miṭʰ-il  | mis-il ~ mičʰ-il ~ mi-til  | ‘bottom’  | Kang  |
| b. /cʰ/      | (i) /k’očʰ-il/ | k’očʰ-il | k’os-il ~ k’otʰ-il        | ‘flower’  | Choe  |
|              | (ii) /k’očʰ-a/ | k’očʰ-a  | k’os-a ~ k’otʰ-a ~ k’oč-a | ‘flower’  | Kim, Choe |
Some additional observations can be made from the survey results. The relative preference among the variants may differ depending on the stem and the suffix, as can be seen in (5). For instance, speakers prefer different stem-final variants for different stems even when they are combined with the same suffix -il. [tʰ] is the most preferred in two stems (/kjɔtʰ/, /mɔli-matʰ/), [cʰ] is most preferred in one stem (/k’itʰ/), and [s] is most preferred in the remaining one stem (/sutʰ/). For the same stem /k’itʰ/, [tʰ] is preferred by 89% of the subjects when the stem combines with the locative case marker -e, but only by 24.5% when the stem combines with the accusative marker -il.

(5) The percentage of coronal obstruents in the survey on the preferred stem-final obstruent when the underlying stem-final obstruent is /tʰ/ (chosen from Choe (2004))

<table>
<thead>
<tr>
<th>UR</th>
<th>variants</th>
<th>(i) -i</th>
<th>(ii) -il</th>
<th>(iii) -e</th>
</tr>
</thead>
<tbody>
<tr>
<td>/kjɔtʰ/</td>
<td>tʰ</td>
<td>55.5</td>
<td>96.85</td>
<td></td>
</tr>
<tr>
<td></td>
<td>s</td>
<td>9.05</td>
<td>6.95</td>
<td>2.2</td>
</tr>
<tr>
<td></td>
<td>cʰ</td>
<td>88.5</td>
<td>37</td>
<td></td>
</tr>
<tr>
<td>/k’itʰ/</td>
<td>tʰ</td>
<td>24.5</td>
<td>89</td>
<td></td>
</tr>
<tr>
<td></td>
<td>s</td>
<td>22.4</td>
<td>22.3</td>
<td>10.45</td>
</tr>
<tr>
<td></td>
<td>cʰ</td>
<td>77.3</td>
<td>53.05</td>
<td></td>
</tr>
<tr>
<td>/mɔli-matʰ/</td>
<td>tʰ</td>
<td>55.45</td>
<td>93.8</td>
<td></td>
</tr>
<tr>
<td></td>
<td>s</td>
<td>20.55</td>
<td>10.05</td>
<td>5.85</td>
</tr>
<tr>
<td></td>
<td>cʰ</td>
<td>78.35</td>
<td>34.1</td>
<td></td>
</tr>
<tr>
<td>/sutʰ/</td>
<td>tʰ</td>
<td>21.85</td>
<td>47.5</td>
<td></td>
</tr>
<tr>
<td></td>
<td>s</td>
<td>50.55</td>
<td>52.4</td>
<td>49.6</td>
</tr>
<tr>
<td></td>
<td>cʰ</td>
<td>48.8</td>
<td>24.45</td>
<td></td>
</tr>
</tbody>
</table>

(Suffixes: /-i/ = nominative, /-il/ = accusative, /-ilo/ = locative/instrumental, /-in/ = topic, /-el/ = dative/locative, /-a/ = vocative)
There are some tendencies in the data. First, [s] is, in general, the most favored stem-final variant. Sometimes, it is favored even over standard ones. Results of a survey by E. Kang et al. (2004) show that in the choice of the stem-final obstruent of /mokčač/-in/ ‘uvula, topicalized’, [s] is preferred by 81.2% of 71 Seoul Korean subjects, aged 20-39 years, whereas only 18.8% of the subjects prefer the underlying [č]. Second, [cʰ] is favored when the stem combines with [i]-initial suffixes such as -il, -in and -ilo, as stated by E. Kang et al. (2004: 12) and Choe (2004: 21). Third, as can be seen in (6), [tʰ] is favored when the stem combines with [e]-initial suffixes such as -e ‘locative/dative’ and -esə ‘locative’ (Ko 1989; Y. Kang 2003a, b; Choe 2004; E. Kang et al. 2004). So, the relative preference of [cʰ] over [tʰ] before [i]-initial suffixes is reversed in the case of [e]-initial suffixes. Finally, [č] and [t] are rarely adopted as a variant.

(6) Average ratio of standard pronunciation for 15 nouns with stem-final [tʰ] (from Choe (2004: 23, Table 20))

<table>
<thead>
<tr>
<th>Suffix types</th>
<th>standard pronunciation</th>
<th>percentage of standard pronunciation</th>
</tr>
</thead>
<tbody>
<tr>
<td>i</td>
<td>cʰ</td>
<td>60.82</td>
</tr>
<tr>
<td>e(sə)</td>
<td>tʰ</td>
<td>76.04</td>
</tr>
<tr>
<td>ilo, il</td>
<td>tʰ</td>
<td>50.76</td>
</tr>
</tbody>
</table>

The survey results also confirm the observation reported in the previous studies, mentioned above, that noncoronal aspirated/tense obstruents are in free variation with their lenis counterparts when the stems are combined with vowel-initial suffixes, as shown below:

(7) Stem-final obstruents = Noncoronal (Kim (2003), Choe (2004), E. Kang et al. (2004) and others)

<table>
<thead>
<tr>
<th>Stem-final C</th>
<th>UR (= standard)</th>
<th>Variants</th>
<th>gloss</th>
</tr>
</thead>
<tbody>
<tr>
<td>/pʰ/</td>
<td>pʰ-e</td>
<td>ipʰ-e</td>
<td>‘leaf’</td>
</tr>
<tr>
<td></td>
<td>supʰ-i</td>
<td>sup-i</td>
<td>‘woods’</td>
</tr>
<tr>
<td>/kʰ/</td>
<td>puakʰ-e</td>
<td>puak-e</td>
<td>‘kitchen’</td>
</tr>
<tr>
<td></td>
<td>sepkjoknjkʰ-e</td>
<td>sepkjoknjk-e</td>
<td>‘dawn’</td>
</tr>
</tbody>
</table>

Speakers differ in the preference of final noncoronal obstruent variants mainly depending on the stem, as illustrated by the following table:
(8) The percentage of noncoronal obstruents in a survey on the preferred stem-final obstruent (from Choe (2004)):

<table>
<thead>
<tr>
<th></th>
<th>( p^h )</th>
<th>( p )</th>
<th>( k^h )</th>
<th>( k )</th>
</tr>
</thead>
<tbody>
<tr>
<td>jəph ‘side’</td>
<td>90.49</td>
<td>9.51</td>
<td>puəkk ‘kitchen’</td>
<td>31.02</td>
</tr>
<tr>
<td>ap’h ‘front’</td>
<td>89.88</td>
<td>10.12</td>
<td>səpəknək ‘dawn’</td>
<td>34.52</td>
</tr>
<tr>
<td>nip’h ‘swarm’</td>
<td>82.88</td>
<td>17.12</td>
<td></td>
<td></td>
</tr>
<tr>
<td>sup’h ‘woods’</td>
<td>79.4</td>
<td>20.6</td>
<td></td>
<td></td>
</tr>
<tr>
<td>namuipl’h ‘leaf’</td>
<td>78.71</td>
<td>21.29</td>
<td></td>
<td></td>
</tr>
<tr>
<td>səp’h ‘prop’</td>
<td>59.94</td>
<td>40.06</td>
<td></td>
<td></td>
</tr>
<tr>
<td>pjačip’h ‘rice straw’</td>
<td>43.37</td>
<td>56.63</td>
<td></td>
<td></td>
</tr>
<tr>
<td>mulipl’h ‘knee’</td>
<td>35.35</td>
<td>64.65</td>
<td></td>
<td></td>
</tr>
<tr>
<td>očilapl’h ‘the front of an outer garment’</td>
<td>29.64</td>
<td>70.36</td>
<td></td>
<td></td>
</tr>
<tr>
<td>həŋkəpl’h ‘patch’</td>
<td>18.83</td>
<td>81.17</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>60.84</td>
<td>39.15</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

All the observations about stem-final obstruent variations, discussed thus far, can be summarized as follows:

(9) A summary of observations about stem-final obstruent variation in nouns

a. Five obstruents \([s, t^h, č^h, č, t]\) are involved in free variation of stem-final coronal obstruents.

b. \([s]\) is, in general, the most favored variant.

c. \([č^h]\) is favored when the stem combines with \([i]\)-initial suffixes.

d. \([t^h]\) is favored when the stem combines with \([e]\)-initial suffixes.

e. \([č]\) and \([t]\) are least favored variants.

f. Noncoronal aspirated/tense obstruents are in free variation with their lenis counterparts.

Before turning to the analysis of these observations, it should be clarified how I interpret the data of the survey results. In Choe, subjects were asked about what sounds they actually produce in stem-final position of each target word whereas in E. Kang et al., subjects were asked about what sounds they prefer in the same condition. In both surveys, the subjects were instructed to choose only one sound for each target word. Different subjects may choose different sounds. Thus, the data in (4) actually shows the list of speaker-dependent variants. But since the same speaker may use different variant forms at different times, the sounds chosen by each subject in both surveys must be her/his most preferred ones among free variants that s/he would produce in natural
speech. Thus, I assume that the survey data like the one in (4) may indicate the range of free variants adopted by an average speaker in natural speech. Similarly, the data like the one in (5), reporting the relative number of speakers choosing different variants for each target word, may indicate an average speaker’s relative preference or acceptability among free variants of noun stems. (See Y. Kang (2003a,b) for a discussion of the relative acceptability of different stem-final variants of certain Korean nouns which differ in token frequency.)

2. Frequency facts
In this section, I will discuss frequency facts of Korean nouns, showing that all the observations about noun stem-final variations, summarized in (9), are consistent with the distribution of final obstruents of suffixed nouns. Based on the discussion, I will claim that the distribution of final obstruents determines Korean speakers’ relative preference of variants. I will then discuss the previous accounts of some relevant facts, pointing out their problems.

2.1 Distribution of final obstruents
The distribution of final obstruents of Korean nouns is shown in (10), which is based on the text corpus of Sejong Project, established from 1999 to 2001. The distribution data in (10b), calculated from Kang and Kim (2004) employing the 5.5 million word corpus as a database, can be considered as an updated version of that in (10a), which Albright (2005a) calculated from Kim and Kang (2000), employing the 1.5 million word corpus database.¹

(10) Distribution of final obstruents of nouns

<table>
<thead>
<tr>
<th></th>
<th>Labial</th>
<th>Coronal</th>
<th>Velars</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>p</td>
<td>t</td>
<td>k</td>
</tr>
<tr>
<td>p¹</td>
<td>1360</td>
<td>113</td>
<td>5994</td>
</tr>
<tr>
<td>pʰ</td>
<td>64</td>
<td>0</td>
<td>18</td>
</tr>
<tr>
<td>s</td>
<td>375</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

¹ Kang and Kim’s database (2004) includes not only mono-morphemic nouns but also compound nouns which are listed as separate entries in Standard Korean Dictionary (written in Korean [phyocwunkwaketaesaecon], published in 1999, Dusan Dong-A). For the distribution in (10b), I have removed English loanwords from Kang and Kim’s database by hand-checking.

<table>
<thead>
<tr>
<th>Labial</th>
<th>Coronal</th>
<th>Velars</th>
</tr>
</thead>
<tbody>
<tr>
<td>p</td>
<td>2370</td>
<td>t</td>
</tr>
<tr>
<td>pʰ</td>
<td>81</td>
<td>tʰ</td>
</tr>
<tr>
<td>p'</td>
<td>0</td>
<td>t'</td>
</tr>
<tr>
<td>ċ</td>
<td></td>
<td>ċʰ</td>
</tr>
<tr>
<td>ċ'</td>
<td></td>
<td></td>
</tr>
<tr>
<td>s</td>
<td></td>
<td>sʰ</td>
</tr>
<tr>
<td>s'</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

From the frequency data in (10), the following hierarchies among final obstruents may fall out:

(11) Frequency hierarchies among final obstruents

a. Coronal: \( s \gg ċʰ \gg tʰ \gg ċ, t, (t', ċ', s') \)

b. Noncoronal: \( k \gg kʰ, k' \)

\( p \gg pʰ \)

Notice that these frequency hierarchies are consistent with observations on noun-final obstruent variations, discussed above and summarized in (9). The highest preference of [s]-final variant forms is matched by the highest type frequency of [s]-final noun stems. The lowest preference of [č/t]-final variant forms is matched by the lowest frequency of [č/t]-final stems. This frequency matching may indicate that the relative frequency of final obstruents in the lexicon determines their relative preference in the choice of final variants in free variation (cf. Albright 2005a). A similar frequency matching can be conceived for the variants ending in noncoronal lenis stops. The frequent occurrence of lenis-final variant forms for stems ending in underlying noncoronal aspirated/tense stops may be motivated by the higher lexical frequency of the noun stems ending in noncoronal lenis stops /p, k/ compared to those ending in aspirated/tense stops, /pʰ, kʰ, k’/.

To check whether the frequency also determines the remaining observations (9c,d), I calculated type frequencies of final obstruents of suffixed nouns from the 5.5 million word text corpus of Sejong Project.² (12) shows the distribution of final coronal obstruents of noun stems when they are combined with four different suffixes, two [i]-initial and two [e]-initial.

The numbers in the table in (12) indicate how many noun stems ending with the

² As in (10b), the database for frequency count includes compound nouns, which are listed as separate entries in Standard Korean Dictionary, and has been hand-checked to remove English loanwords.
corresponding obstruent are inflected with the corresponding suffix. For instance, the corpus contains 149 different nouns with the suffix \(-il\) whose stems end with \([s]\) as in /nas-il/ ‘sickle, accusative’. The order of frequency (from most to least) is \(s >> ch^h >> t^b >> č >> t\) before [i]-initial suffixes, but \(s, t^b >> ch^h >> č >> t\) before [e]-initial suffixes. Notice that the relative frequency between \([č]\) and \([t]\) is reversed where the relevant cells are shaded. This frequency reversal matches with the preference reversal of the two variants depending on the suffix type, discussed in the previous section and summarized in (9c,d).

(12) Distribution of final coronal obstruents of suffixed nouns

<table>
<thead>
<tr>
<th>Final C (↓) vs. suffix (→)</th>
<th>il</th>
<th>in</th>
<th>e</th>
<th>esω</th>
</tr>
</thead>
<tbody>
<tr>
<td>(s)</td>
<td>149</td>
<td>77</td>
<td>57</td>
<td>28</td>
</tr>
<tr>
<td>(t^b)</td>
<td>42</td>
<td>19</td>
<td>56</td>
<td>32</td>
</tr>
<tr>
<td>(č^h)</td>
<td>67</td>
<td>36</td>
<td>21</td>
<td>8</td>
</tr>
<tr>
<td>(č)</td>
<td>7</td>
<td>7</td>
<td>6</td>
<td>1</td>
</tr>
<tr>
<td>(t)</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

Consequently, the observed suffix-specific difference in the preference between the two coronal variants, \([č^h]\) and \([t^b]\), is attributable to the suffix-specific asymmetry in the type frequency of noun stems ending in these obstruents. Notice that other observations about stem-final obstruent variations summarized in (9) are also consistent with the distribution of final obstruents of suffixed nouns in (12). Overall, \([s]\)-final stems are the most frequent whereas \([č]\)-final and \([t]\)-final stems are least frequent. All this suggests that the relative frequency of inflected forms, not bare stems, determines the relative preference of coronal variants.

The table in (13) shows the distribution of final noncoronal obstruents of noun stems when they are combined with the four suffixes, \(-il\), \(-in\), \(-e\), and \(-esω\). The relative frequency here is not different from the one shown in (11b). Regardless of the suffixes attached, the number of stems ending with lenis stops far exceeds the number of those ending with corresponding aspirated (and tense) stops.

(13) Distribution of final noncoronal obstruents of suffixed nouns

<table>
<thead>
<tr>
<th>Final C (↓) vs. suffix (→)</th>
<th>il</th>
<th>in</th>
<th>e</th>
<th>esω</th>
</tr>
</thead>
<tbody>
<tr>
<td>(p)</td>
<td>720</td>
<td>428</td>
<td>522</td>
<td>223</td>
</tr>
<tr>
<td>(p^h)</td>
<td>34</td>
<td>15</td>
<td>23</td>
<td>15</td>
</tr>
<tr>
<td>(k)</td>
<td>2472</td>
<td>1382</td>
<td>1485</td>
<td>687</td>
</tr>
<tr>
<td>(k^h)</td>
<td>3</td>
<td>4</td>
<td>12</td>
<td>8</td>
</tr>
<tr>
<td>(k')</td>
<td>2</td>
<td>2</td>
<td>6</td>
<td>3</td>
</tr>
</tbody>
</table>
The frequency facts of suffixed nouns, presented above, are summarized as below:

(14) A summary of crucial frequency facts

a. Among stem-final coronal obstruents, [s] in general has highest frequency, and [č, t] have lowest frequency, regardless of the following suffix.

b. Before [e]-initial suffixes, [tʰ] has a frequency higher than [čʰ].

c. Before [i]-initial suffixes, [čʰ] is higher in frequency than [tʰ].

d. Among noncoronal obstruents, lenis stops have a higher frequency than their aspirated/tense counterparts.

From the finding that these frequency facts are consistent with patterns of stem-final obstruent variation, it may be concluded that the relative frequency of nouns ending in a certain obstruent determines the degree of its preference in the choice of the stem-final variant of nouns. Moreover, type frequency of suffixed forms, not bare stems, is the determining factor in the preference of variants since suffix-specific variation patterns can be explained only in terms of the frequency of suffixed forms, not bare stems.

2.2 Other explanations of suffix-specific preferences of variants

It has been observed in previous research on Korean phonology and morphology that there are two exceptional cases in which the coronal stem-final variant [s] is rarely adopted. First, [čʰ]-final variants frequently occur before [i]-initial suffixes. Second, [tʰ]-final stems may often surface as such before [e]-initial suffixes. Notice that these two exceptional cases have been discussed and attributed, in the previous section, to the suffix-specific asymmetry in the relative frequency of noun stems ending in different coronal obstruents. I will now discuss other explanations which have been proposed in the literature, pointing out serious problems with them. Most previous approaches adopt separate mechanisms for the explanation of the two exceptional cases, which itself should be considered disadvantageous.

Let us first consider previous accounts of the frequent occurrence of [čʰ]-final variants before [i]-initial suffixes. Its standard explanation (for instance, Y. Kang 2003b; Choe 2004) relies on Paradigm Uniformity effects. Underlyingly /tʰ/-final stems are realized as allomorphs ending with [čʰ] in the palatalizing context, i.e., before [i]-initial suffixes, and these allomorphs influence the realization of allomorphs occurring in the non-palatalizing context, here before [i]-initial suffixes. For instance, under the influence of the standard output form [pačʰi] of /patʰ-i/ ‘field, nominative’, /patʰ-il/ ‘field, accusative’ is realized as [pačʰi]. There are many problems with this account. First, recall that all five coronal obstruent variants may occur when the stems are combined with various different vowel-initial suffixes, including /i/-initial suffixes, although the relative preference among them may differ. If [čʰ]-final forms occur due to
Paradigm Uniformity, this explanation cannot be extended to the explanation of the other variants, and thus separate mechanisms are needed to explain the occurrence of the other variants including [c] and [t]-final forms. Second, as will be discussed below, most previous studies on Korean stem-final variations assume that isolation forms are the base in explaining the variations, mainly based on the fact that nominative and accusative case markers are often omitted in natural speech, and thus the isolation forms have the highest token frequency. There is no comparable justification for claiming that the base of Paradigm Uniformity is the nominative forms, inflected with the palatalizing suffix –i. If it is the case that Paradigm Uniformity is operative, depalatalization would be more likely to occur under the general assumption that isolation forms are the base.

Third, section 2.1 shows that [c]-final stems have relatively high frequency in the corpus, only when they are combined with [i]-initial suffixes. The fact that [c]-final variants occur frequently in the same context, i.e., before [i]-initial suffixes, would be accidental within previous Paradigm Uniformity-based accounts. Finally and most importantly, [c]-final variants may occur even when the underlying obstruent is not /t/: for example, [nac-h-lo] for /nac-ilo/. Notice that in its standard pronunciation, the underlying stem-final obstruent /c/ should surface as [c], not [c], in the palatalizing context, and thus the occurrence of [c] before a suffix -ilo cannot be due to the palatalized members of a paradigm. In conclusion, Paradigm Uniformity-based accounts cannot correctly explain the frequent occurrence of [c]-final variants before [i]-initial suffixes.

We are now in a position to consider the previous accounts of the frequent occurrence of [t]-final forms before [e]-initial suffixes. In the previous studies (Y. Kang 2003b; Choe 2004 and others), this suffix-specific fact is attributed to the high token frequency of the locative suffix –e which is never omitted even in conversational speech unlike suffixes such as –i ‘nominative’ and -il ‘accusative’. It has been claimed that high token frequency of noun forms suffixed with -e helps [t]-final stems to resist an analogical change to [s]-final forms or Paradigm Uniformity, and thus [t]-final stems are likely to surface as such before the suffix -e. This proposal is also subject to some difficult problems. First, if the frequent occurrence of [t]-final forms is attributable to the frequent retention of the suffix –e, not only underlying /t/ but also other underlying obstruents such as /c/ should occur significantly more often in combination with –e. This is not true with /c/-final nouns, as shown below.

(15) Average ratio of standard pronunciation of 11 / c/-final nouns, classified by suffix types (Choe 2004: 67, Table 70)

<table>
<thead>
<tr>
<th>Suffix types</th>
<th>percentage of [c]</th>
</tr>
</thead>
<tbody>
<tr>
<td>i</td>
<td>61.92</td>
</tr>
<tr>
<td>e(sa)</td>
<td>48.45</td>
</tr>
<tr>
<td>ilo, il</td>
<td>55.66</td>
</tr>
</tbody>
</table>
In addition, the occurrence of stem-final [tʰ] is not confined to underlyingly [tʰ]-final stems. For instance, [tʰ]-final variants are often attested with noun stems ending in /cʰ/, especially when followed by locative suffix –e(s): /sucʰ-čel/ [suhčel] ‘charcoal’ and /stalkačʰ-čel/ [stalaččel] ‘the skin (surface)’. Therefore, the previous account relying on analogy-resistance cannot be considered as a general solution of the issue involving the frequent occurrence of [tʰ]-final variants before [e]-initial suffixes.

In conclusion, only the frequency-based account can provide a unified and plausible explanation of the suffix-specific variation patterns involving [cʰ] and [tʰ]-final forms. In the following section, I will explore the stem-final variations of Korean bound stems which typically combine with the verb –ha ‘do’. Specifically, I will investigate whether the same type of frequency effect observed in noun variations can be seen in these other variations, in order to determine an optimal analysis of the variation in Korean stem-final obstruents.

3. Variation in Bound stems

In Korean, certain noun stems are combined with the verb stem ha ‘do’ to form compound verbs, as shown in (16a). This compounding is productive, and it is a usual way to derive verbs from loanwords, as illustrated in (16b).

(16) Compound verbs

a. Noun stem + ha = verb stem (Sohn 2001)
   (i) il     ‘work, noun’
   (ii) il-ha-ta ‘work, verb, infinitive’
   (iii) il-ha-ko ‘work, verb, connective’

b. Loanword + ha = verb stem
   (i) keim    ‘game, noun’
   (ii) keim-ha-ta ‘play a game, infinitive’
   (iii) keim-ha-ko ‘play a game, connective’

The same type of compounding is also adopted to derive verbs from a certain type of bound stems, as shown below:3

(17) Compound verbs: Bound stem + ha = verb stem

   (i) pʰokin-ha-ta ‘warm, infinitive’
   (ii) pʰokin-ha-ko ‘warm, connective’
   (iii) pʰisin-ha-ta ‘similar, infinitive’
   (iv) pʰisin-ha-ko ‘similar, connective’
   (v) taptap-ha-ta ‘stuffy, infinitive’
   (vi) taptap-ha-ko ‘stuffy, connective’

3 The resulting compounds would be more correctly classified as adjectives. But, in Korean, the distinction between verbs and adjectives is not always clear mainly because they are inflected in a similar way. Since the distinction is not relevant to any important aspects of this paper, for simplicity’s sake I will consider the compounds as verbs in this paper.

13
As the name indicates, these bound stems normally do not stand by themselves although some of them may combine with the adverbializing suffix –i in addition to the verb ha: for example, pʰok-i ‘warmly’. Notice that most bound stems of this type are mimetic vocabulary. As shown in (17iii-v), many of them in fact have a form of reduplication which is typical of mimetic and onomatopoeic words in Korean. Considering such word forms and meanings, the bound stems under consideration may belong to a type of adverb. What is important here is that they cannot be classified as nouns. Due to these special properties, they are sometimes classified as an independent category in Korean morphology (for instance, Kang and Kim 2004).

The focus of the present study is on the topicalized forms of these bound stems in which a topic marker -in is inserted between the bound stem and ha, as shown in (18). When the bound stems ending with obstruents are topicalized, as in (18ii-v), the final obstruents appear in prevocalic position, which is a probable location for free variation.

(18) Topic formation of bound stems: [stem + in + ha]

(i) pʰokin-in-ha-ta ‘warm, topicalized, infinitive’
(ii) pʰis-i-in-ha-ta ‘similar, topicalized, infinitive’
(iii) taptap-in-ha-ta ‘stuffy, topicalized, infinitive’
(iv) t’okt’ok-in-ha-ta ‘clever, topicalized, infinitive’
(v) t’aкт’ak-in-ha-ta ‘hard, topicalized, infinitive’

To explore the phonetic realizations of the final obstruents, I consulted with native Korean speakers. The next section discusses the process and method of the judgment survey. For now, I want to emphasize that although the topic formation of bound stems is completely acceptable to native Korean speakers, it is very rarely used in actual speech. So, the judgment task, described below, would be a type of Wug-test (Burko 1958) for most native Korean speakers who participated in the experiment. (See section 4 for a discussion of this assumption.)

3.1 Judgment Survey

Bound stems employed in the task are listed in (19), where the stems are transcribed based on standard Korean orthography.
(19) Test stems

<table>
<thead>
<tr>
<th>place of stem-final obstruent</th>
<th>Stems</th>
<th>glosses</th>
<th>Stems</th>
<th>glosses</th>
</tr>
</thead>
<tbody>
<tr>
<td>coronal (11 stems)</td>
<td>pisis</td>
<td>similar</td>
<td>k’ek’is</td>
<td>clean</td>
</tr>
<tr>
<td></td>
<td>t’at’is</td>
<td>warm</td>
<td>t’ət’as</td>
<td>aboveboard</td>
</tr>
<tr>
<td></td>
<td>himus</td>
<td>pleased</td>
<td>nikis</td>
<td>relaxed</td>
</tr>
<tr>
<td></td>
<td>iičas</td>
<td>dignified</td>
<td>e’t’is</td>
<td>pitiful</td>
</tr>
<tr>
<td></td>
<td>p’utis</td>
<td>satisfied</td>
<td>č’alis</td>
<td>stimulating</td>
</tr>
<tr>
<td></td>
<td>pantis</td>
<td>decent</td>
<td></td>
<td></td>
</tr>
<tr>
<td>labial (11 stems)</td>
<td>pokčap</td>
<td>crowded</td>
<td>takip</td>
<td>imminent</td>
</tr>
<tr>
<td></td>
<td>səŋkip</td>
<td>impatient</td>
<td>pikap</td>
<td>cowardly</td>
</tr>
<tr>
<td></td>
<td>čokip</td>
<td>hasty</td>
<td>mihip</td>
<td>insufficient</td>
</tr>
<tr>
<td></td>
<td>taptap</td>
<td>stuffy</td>
<td>səpsəp</td>
<td>sorry</td>
</tr>
<tr>
<td></td>
<td>k’amučapčap</td>
<td>darkish</td>
<td>kapkap</td>
<td>stuffy</td>
</tr>
<tr>
<td></td>
<td>č’ipč’ip</td>
<td>unsatisfied</td>
<td></td>
<td></td>
</tr>
<tr>
<td>velar (12 stems)</td>
<td>simkak</td>
<td>serious</td>
<td>palamčik</td>
<td>desirable</td>
</tr>
<tr>
<td></td>
<td>solčik</td>
<td>honest</td>
<td>naknək</td>
<td>enough</td>
</tr>
<tr>
<td></td>
<td>t’aťąk’ak</td>
<td>hard</td>
<td>t’əkt’ok</td>
<td>clever</td>
</tr>
<tr>
<td></td>
<td>k’imčik</td>
<td>horrible</td>
<td>kipkjak</td>
<td>drastic</td>
</tr>
<tr>
<td></td>
<td>makmak</td>
<td>desolate</td>
<td>t’ətils’ək</td>
<td>noisy</td>
</tr>
<tr>
<td></td>
<td>katik</td>
<td>full</td>
<td>s’iks’ik</td>
<td>lively</td>
</tr>
</tbody>
</table>

These stems were combined with a topic marker -in followed by an inflected form of the verb *ha*. The resulting target words, together with their possible phonetic forms, were put in the context of sentences in the format shown in (20). The phonetic forms of each target word differ only in the stem-final obstruent. Five coronal obstruents [s, tʰ, čʰ, č, t] and two noncoronal ones, lenis and aspirated, were employed as the stem-final obstruent.

(20) Example experimental sentences

a. Final obstruent = coronal

\[
\text{wancənhi katʰci-nin anh-țəlato pisis-in he-jači. (ha + őjači } \rightarrow \text{ hejači)}
\]

\[
\begin{align*}
\text{[pisisin]} & \quad ( ) \\
\text{[pisiťʰin]} & \quad ( ) \\
\text{[pisičʰin]} & \quad ( ) \\
\text{[pisičin]} & \quad ( ) \\
\text{[pisiťin]} & \quad ( )
\end{align*}
\]

‘completely’ ‘same-Top’ ‘not-although’ ‘similar-Top’ ‘do-should’
‘Even if (they) are not completely identical, they should be similar.’

b. Final obstruent = labial

ki salam sォkjョk-i sォkip-in ha-nte, kileto もًakhe.

[sォkipin] (     )

[sォkip^in] (     )

‘the’ ‘person’ ‘personality-Nom’ ‘impatient-Top’ ‘do-though’ ‘nonetheless’ ‘nice’ ‘Although his personality is impatient, he is nice.’

The experimental sentences were randomly ordered and written on A4 size paper in Korean orthography which is mostly phonemic (See Appendix for a complete list of the experimental sentences). Thirty paid Seoul Korean speakers, aged 20s to 40s, were asked to evaluate their degree of acceptability by rating 1 through 4 in the order of acceptability: 4 (good), 3 (O.K.), 2 (not impossible), 1 (impossible). I will call these numbers 4-1 “acceptability scores”. The subjects were allowed not to respond in case the sentences sound too awkward for them to determine their acceptability.

3.2 Results

I report the results with a ‘degree of acceptability’ index, defined below.

(21) Degree of Acceptability

** At least half the subjects (i.e. 15) choose 4 or 3 (meaning at least “O.K.”)

* At least half the subjects (i.e. 15) choose 4 or 3 or 2 (meaning at least “not impossible”)

x otherwise (meaning “impossible”)

To illustrate how to assign these indexes, consider the table in (22) for responses to the variants of the test stem /pisι_-in/ ‘similar, topic’. The numbers shown under each acceptability score indicate the number of subjects who chose the corresponding score for the corresponding variant. Degree of acceptability is determined based on the numbers provided under “4+3” and “4+3+2”. Notice, for instance, that ** is assigned to the variant [pisisin] since the number under “4+3”, i.e., 28, is bigger than 15. In contrast, * is assigned to the variant [pisit^in] since the number under “4+3”, i.e., 9, is smaller than 15 and the number under “4+3+2”, i.e., 20 is bigger than 15. The mean acceptability score for each variant, shown on the last column of the table in (22), is also calculated.
Let us first consider the data of target stems ending with the coronal obstruent. The table in (23) shows degrees of acceptability as well as mean acceptability scores of target variants for all test stems ending with the coronal obstruent. According to the degree of acceptability index, all five variants are mostly acceptable, i.e., at least “not impossible” (46 out of 55 = 84%) although the exact degrees differ depending on the stem and variant. [s]-final variant is most acceptable in all test stems. [čʰ]-final variant is at least “not impossible” in all test stems and it is the second best choice. The remaining three variants ending with [tʰ, č, t] are “impossible”, marked with shading, at least for some test stems. The relative acceptability among these three variants can also be seen from the mean acceptability scores in (23b). It seems [tʰ]-final forms are more acceptable than the other two, but the difference between [č] and [t] is negligibly small. Thus, the overall order of acceptability (from most to least) is s >> čʰ >> tʰ >> č, t.

(23) Acceptability of variants (Stem-final C = Coronal)

<table>
<thead>
<tr>
<th></th>
<th></th>
<th>s</th>
<th>tʰ</th>
<th>čʰ</th>
<th>č</th>
<th>t</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. pis</td>
<td>**</td>
<td>**</td>
<td>*</td>
<td>*</td>
<td>x</td>
<td>*</td>
</tr>
<tr>
<td>2. k’eš</td>
<td>**</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>3. t’aš</td>
<td>**</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>4. t’os</td>
<td>**</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>5. himuš</td>
<td>**</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>x</td>
<td></td>
</tr>
<tr>
<td>6. nikiš</td>
<td>**</td>
<td>*</td>
<td>*</td>
<td>x</td>
<td></td>
<td></td>
</tr>
<tr>
<td>7. ičos</td>
<td>**</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>8. etš</td>
<td>**</td>
<td>x</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td></td>
</tr>
<tr>
<td>9. p’utis</td>
<td>**</td>
<td>x</td>
<td>*</td>
<td>x</td>
<td>*</td>
<td>x</td>
</tr>
<tr>
<td>10. č’alis</td>
<td>**</td>
<td>*</td>
<td>*</td>
<td>x</td>
<td>x</td>
<td></td>
</tr>
<tr>
<td>11. pantis</td>
<td>**</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td></td>
</tr>
</tbody>
</table>
b. Mean acceptibility scores

<table>
<thead>
<tr>
<th>stem-final obstruent</th>
<th>s</th>
<th>t&lt;sup&gt;h&lt;/sup&gt;</th>
<th>Č&lt;sup&gt;h&lt;/sup&gt;</th>
<th>Č</th>
<th>t</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>3.78</td>
<td>2.03</td>
<td>2.39</td>
<td>1.92</td>
<td>1.90</td>
</tr>
<tr>
<td>standard deviation</td>
<td>0.35</td>
<td>0.57</td>
<td>0.59</td>
<td>0.72</td>
<td>0.61</td>
</tr>
</tbody>
</table>

Notice that this relative acceptability among variants of topicalized bound stems is mostly consistent with variation patterns of noun stems and the distribution of final obstruents of noun stems inflected with a topic marker –<i>in</i>, repeated below:

(24) Distribution of final obstruents of topicalized noun stems ending with coronals
(topic marker = in)

<table>
<thead>
<tr>
<th>stem-final obstruent</th>
<th>s</th>
<th>Č&lt;sup&gt;h&lt;/sup&gt;</th>
<th>t&lt;sup&gt;h&lt;/sup&gt;</th>
<th>Č</th>
<th>t</th>
</tr>
</thead>
<tbody>
<tr>
<td>type frequency</td>
<td>77</td>
<td>36</td>
<td>19</td>
<td>7</td>
<td>0</td>
</tr>
</tbody>
</table>

Similar interesting observations can be made with the data of bound stems ending with noncoronal obstruents. The table in (25) shows degrees of acceptability and mean acceptibility scores of target variants for test stems ending with the noncoronal obstruent. As above, multiple forms are acceptable. In all cases, more than one form are at least “not impossible”. According to a “degree of acceptability” index, lenis-final variants are very much acceptable for both sets of stems ending with labial and velar stops. Among less acceptable variants ending with aspirated stops, the labial aspirated stop is more acceptable for the stem-final variant than the velar since the aspirated labial variants are “not impossible” for all eleven labial-final test stems whereas the aspirated velar ones are “impossible” for five out of twelve velar-final test stems. This relative acceptability can be confirmed by the mean acceptibility scores in (25b).

(25) Acceptability of variants (Stem-final C = Noncoronal)

a. Degree of acceptability

<table>
<thead>
<tr>
<th></th>
<th>p</th>
<th>p&lt;sup&gt;h&lt;/sup&gt;</th>
<th>Velar</th>
<th>k</th>
<th>k&lt;sup&gt;h&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>pok čap</td>
<td>** *</td>
<td></td>
<td>simkak</td>
<td>** x</td>
<td></td>
</tr>
<tr>
<td>takip</td>
<td>** *</td>
<td></td>
<td>palamčik</td>
<td>** *</td>
<td></td>
</tr>
<tr>
<td>soŋkip</td>
<td>** *</td>
<td></td>
<td>solčik</td>
<td>** *</td>
<td></td>
</tr>
<tr>
<td>pikap</td>
<td>** *</td>
<td></td>
<td>naknok</td>
<td>** *</td>
<td></td>
</tr>
<tr>
<td>čokip</td>
<td>** *</td>
<td></td>
<td>t’akt’ak</td>
<td>** x</td>
<td></td>
</tr>
<tr>
<td>mihip</td>
<td>** *</td>
<td></td>
<td>t’okt’ok</td>
<td>** *</td>
<td></td>
</tr>
<tr>
<td>taptap</td>
<td>** *</td>
<td></td>
<td>k’ime’ik</td>
<td>** x</td>
<td></td>
</tr>
<tr>
<td>sapsəp</td>
<td>** *</td>
<td></td>
<td>kipkək</td>
<td>** *</td>
<td></td>
</tr>
</tbody>
</table>
b. Mean acceptability score

<table>
<thead>
<tr>
<th>stem-final obstruent</th>
<th>p</th>
<th>pʰ</th>
<th>k</th>
<th>kʰ</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>3.76</td>
<td>2.14</td>
<td>3.85</td>
<td>1.79</td>
</tr>
<tr>
<td>Standard Deviation</td>
<td>0.35</td>
<td>0.77</td>
<td>0.34</td>
<td>0.69</td>
</tr>
</tbody>
</table>

Notice that the high acceptability of lenis-final forms is matched by the high type frequency of topicalized noun stems ending in noncoronal lenis stops, as repeated in (26), and the frequent occurrence of lenis-final variants for the stems ending in noncoronal aspirated/tense stops, discussed earlier.

(26) Distribution of final obstruents of topicalized noun stems ending with noncoronals

<table>
<thead>
<tr>
<th>stem-final obstruent</th>
<th>p</th>
<th>pʰ</th>
<th>k</th>
<th>kʰ</th>
</tr>
</thead>
<tbody>
<tr>
<td>type frequency</td>
<td>428</td>
<td>15</td>
<td>1382</td>
<td>4</td>
</tr>
</tbody>
</table>

As for aspirated variant forms, the higher acceptability of labial-final ones, relative to velar-final ones, is matched by the higher ratio of aspirated stem-final stops among labials (15/443), compared to aspirated ones among velars (4/1386). There is some indication that there might be the comparable difference in the native Korean speakers’ preference between labial-final and velar-final variants of noun stems. The data in (8), which is chosen from Choe (2004), shows that variants ending in the aspirated stop are more likely to be preferred when the stop is labial (60.84%), compared to when it is velar (32.77%). But, the data is not sufficient and, more importantly, imbalanced. Choe’s (2004) survey employed ten test stems ending in a labial stop but only two stems ending in a velar stop. Thus, the final conclusion about the relative preference between labial-final vs. velar-final variants should await further investigation.

To summarize, an important observation, which can be made from the experimental results, just presented, is that Korean speakers’ judgments of acceptability of topicalized bound stems show almost the same type of stem-final obstruent variation as can be seen with nouns. Strikingly, the relative frequency of nouns ending in a certain obstruent determines the relative acceptability of topicalized bound stems ending in this obstruent.

4. Analysis

Let us now discuss the findings of the present study and what can be inferred from them, while seeking an optimal analysis of the stem-final variation patterns in Korean.
First, lexical frequency is matched by the speakers’ preference and acceptability of the stem-final variants. This frequency matching can be understood only under the hypothesis that the speakers are aware of variable lexical patterns, and they behave based on the knowledge (Zuraw 2000, 2005; Bybee 2001; Albright 2002a, b; Albright and Hayes 2003; Pierrehumbert 2003; Hayes and Londe 2006). This hypothesis is strongly supported by the data of the judgment survey with the topic forms of the Korean bound stems, which I assume is a wug-test. As mentioned above, the topic formation of bound stems is legitimate in Korean, and native Korean speakers have no difficulty understanding and producing it. Nevertheless, it is very rarely used in actual speech. I checked the 1.5 million word text corpus of Sejong Project (established in 1999) to obtain the token frequency of topicalized forms of thirty-four test stems employed in the judgment survey. I found only a single instance of a single word [mihip-in] ‘insufficient, topicalized’. The corpus includes no other occurrence of the topicalized forms of the test words despite the fact that all of them frequently occur without the topic marker, immediately followed by ha. Thus, the judgment survey of the present study can be considered as a wug-test although it employs real stems, not made-up stems.

Second, what is matched by the speakers’ preference or acceptability is the type frequency of the suffixed forms, not bare stems. It is then necessary to assume that suffixed forms are listed in the lexicon (Bybee 2001; cf. Zuraw 2002).

Third, multiple variants are attested for the same stem. Maximally five different coronal obstruents may occur as free variants. The coexistence of this many number of variants supports the existence of multiple competing rules or constraints in the mental grammar (Albright 2002a,b, 2005a,b; Albright and Hayes 2003). These rules or constraints should have different applicability since different variants have different preference or acceptability.

Fourth, the relative acceptability of variants of the bound stems is matched by the type frequency of the stems of a different morphological category, i.e., noun. This cross-category matching supports a product-oriented approach, as opposed to a source-oriented one. As discussed by Bybee (2001), in a source-oriented approach, the form of a certain (inflectional) category is derived, in a well-defined way, from the form of

There are several possible answers to a question of why the topic forms of the bound stems are rarely used despite the fact that they are completely acceptable to native Korean speakers. One simple answer might be that situations, in which bound stems need to be topicalized, occur very rarely in actual speech. Another answer, which I think is more plausible, would be the existence of an alternative way to topicalize the bound stems. Topic formation in Korean differs depending on whether the stem is a noun or a verb. Topic markers, -in and -nin, directly attach to the end of noun stems, and the choice depends on whether the stem ends in a consonant or a vowel: for instance, /salam-in/ ‘person, topic’ and /e-a-nin/ ‘car, topic’. In contrast, in the topic formation of the verbs, topic markers attach to the nominal form of verb stems in which a nominalizing suffix –ki attaches to the end of the stem, and then the combined form is followed by the verb ha ‘do’ or re-appearance of the stem. For instance, the topic form of the verb stem /mok/ meaning ‘eat’ is /mok-ki-nin (ha-ta or mok-ta)/. The bound stems under consideration may be topicalized according to the latter formation. The entire compound verb, consisting of the bound stem plus ha, become nominalized through the suffixation of –ki, and then the topic marker is attached to the resulting nominalized form. For instance, the topic form of /pis-ha/ ‘similar’ is / pis-ha-ki-nin (hata)/. This formation seems to be adopted as a default way of topicalizing the bound stems, leading to the rare use of the other topic formation which I adopt for the judgment survey.
some other category. As details will be discussed in section 5, most previous analyses of variation patterns of Korean nouns may be classified as this type of approach since in these approaches, prevocalic allomorphic output forms within a noun paradigm are derived from the isolation form of the same paradigm. In contrast, in a product-oriented approach, a certain lexical item’s output form of a category is determined according to the generalized pattern of the output forms of the same category. Korean variation patterns should be explained in terms of a product-oriented approach since similar variation patterns are observed in two different morphological categories, nouns and bound stems, which have totally different paradigm systems. (See section 5 for an in-depth discussion of this conclusion.)

Based on these findings and related interpretations, I will now provide a possible formal analysis of Korean stem-final variation patterns. The analysis is formalized within the framework of Optimality Theory (OT: Prince and Smolensky 2004), but the standard OT assumptions concerning universal constraints and fixed rankings need to be revised. The native speakers know patterns of the listed suffixed forms, and the knowledge may include phonotactic generalizations on the positioning of segments within words, projecting language-specific constraints (cf. Hayes 1999). The occurrence of Korean variant forms under consideration would be attributable to the following suffix-specific markedness constraints specific to Korean.

(27) Suffix-specific markedness constraints

a. pre-[in] constraints

(i) s/_ in: stem-final coronal obstruent is [s] before a topic marker [in].
(ii) tʰ/_ in: stem-final coronal obstruent is [tʰ] before a topic marker [in].
(iii) ē/h/_ in: stem-final coronal obstruent is [ē/h] before a topic marker [in].
(iv) ē/_ in: stem-final coronal obstruent is [ē] before a topic marker [in].
(v) t/_ in: stem-final coronal obstruent is [t] before a topic marker [in].
(vi) p/_ in: stem-final noncoronal obstruent is [p] before a topic marker [in].
(vii) pʰ/_ in: stem-final noncoronal obstruent is [pʰ] before a topic marker [in].
(viii) k/_ in: stem-final noncoronal obstruent is [k] before a topic marker [in].
(ix) kʰ/_ in: stem-final noncoronal obstruent is [kʰ] before a topic marker [in].

b. pre-[il] constraints

(i) s/_ il: stem-final coronal obstruent is [s] before an accusative marker [il].
(ii) tʰ/_ il: stem-final coronal obstruent is [tʰ] before an accusative marker [il].

... c. pre-[e] constraints

(i) s/_ e: stem-final coronal obstruent is [s] before a locative marker [e].
(ii) tʰ/_ e: stem-final coronal obstruent is [tʰ] before a locative marker [e].

...
d. pre-[esə] constraints

(i) s/_ esə: stem-final coronal obstruent is [s] before a locative marker [esə].
(ii) tʰ/_ esə: stem-final coronal obstruent is [tʰ] before a locative marker [esə]. ... 

These constraints apply regardless of the morphological category of the base, thus capturing the similarity in variation between nouns and bound stems. These constraints are internally ranked, as shown below, based on the type frequency, presented in (12-13) of section 2.1, of relevant suffixed forms in the lexicon.

(28) Rankings of suffix-specific markedness constraints

a. pre-[in] constraints:
   (i) s/_ in >> tʰ/ in >> c/ in, t/_ in
   (ii) p/_ in >> pʰ/ in
   (iii) k/_ in >> kʰ/ in

b. pre-[il] constraints
   (i) s/_ il >> cʰ/ il  >> tʰ/ il >> c/ il, t/_ il
   (ii) p/_ il >> pʰ/ il
   (iii) k/_ il >> kʰ/ il

c. pre-[e] constraints:
   (i) s/_ e, tʰ/ e >> cʰ/ e >> c/ e, t/_ e
   (ii) p/_ e >> pʰ/ e
   (iii) k/_ e >> kʰ/ e

d. pre-[esə] constraints:
   (i) s/_ esə, tʰ/ esə >> cʰ/ esə >> c/ esə, t/_ esə
   (ii) p/_ esə >> pʰ/ esə
   (iii) k/_ esə >> kʰ/ esə

Notice that the proposed rankings may explain the relative preference and acceptability, shown in (9) and (23)/(25), among stem-final variants of nouns and bound stems. For instance, the relative ranking [cʰ/ il >> tʰ/ il] explains the preference of [cʰ]-final noun forms over [tʰ]-final forms before the suffix -il whereas the relative ranking [tʰ/ e >> cʰ/ e] explains the reversed preference before the suffix -e. But, the rankings in (28) should not be fixed since, otherwise, only the most preferred variants would surface due to the top-ranked constraints. In other words, only s-final variants would be attested...
when the stem-final obstruent is coronal. To explain free variation, we may follow the stochastic extension of OT (Boersma and Hayes 2001; Zuraw 2002; Hayes and Londe 2006) in that constraints are ranked in the continuous ranking scale, and each constraint is associated with a range where its position in the scale randomly vary (with a normal distribution), as schematized below:

![Figure 1. Stochastic ranking of pre-[in] constraints](image)

Under the overlap of the associated ranges, constraints with higher ranking values are more frequently, but not always, ranked above those with lower values (For details of this mechanism, see Boersma and Hayes). For instance, s/\_<in would be top-ranked most often, thus producing [s]-final forms most frequently. But, it is less often the case that s/\_<in is ranked low within its range and, at the same time, some other overlapping constraint, say c\_<\_ in, is ranked high, reversing their usual relative ranking (marked with x’s in Figure 1). Then variants like [c_\^]h-final forms would occur. Consequently, the variable ranking model may explain not only the overall relative preference and acceptability of variants but also free variation.

As will be discussed below, the proposed analysis is similar to Albright’s (2005a) account in several respects, one of which is the important role of type frequency in the analysis. But, the crucial difference lies in that my analysis is product-oriented unlike Albright’s source-oriented model.

5. Previous approaches

Most, if not all, previous accounts of stem-final variations of Korean nouns are source-oriented, and concerned mainly with [s]-final variant forms. Let us begin with phonetically-based accounts (for instance, Kim (2001), cf. Kwak (1984)), where s-final forms are derived through an intervocalic (optional) assimilation rule of the type, shown below:

(29) Intervocalic (optional) assimilation: /t, c_h, c/ \rightarrow [s] / V__+V

This approach’s assumption that the process involved is phonetically natural, has been criticized mainly because the rule applies only to noun paradigms (Ko 1989). Also, it says nothing about the occurrence of other coronal variants, not to mention the relative preference and acceptability of the variants. Thus, I reject this approach and it will not be considered again in the remainder of this paper.
All other approaches (Ko 1989; Albright 2002a,b, 2005a,b; Kang 2003a,b; Park 2006) rely on the concept of “analogy” although the formal details may be different. I will first discuss Albright’s Paradigm Learning Model which, I think, is not only most convincing, but also similar to my proposal in many respects.

Albright’s Paradigm Learning Model is a theory about how to learn paradigms during language acquisition. A single most informative form within a paradigm is selected as the Base. Multiple morphological rules, generalized or word-specific, take the Base as the source of the rule, and derive the rest of the paradigm. Example rules of English verb paradigm are shown below:

(30) Example rules for English present-past pairs, the present form = the Base

<table>
<thead>
<tr>
<th>word-specific</th>
<th>generalized</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\emptyset \rightarrow d/#$ plæn __ #</td>
<td>$\emptyset \rightarrow d / __#$</td>
</tr>
<tr>
<td>$i \rightarrow æ / # $fl __ ñ #</td>
<td>$i \rightarrow æ / __ñ$</td>
</tr>
<tr>
<td>$i \rightarrow æ / # $bl __ d #</td>
<td>$i \rightarrow æ / liquid _d$</td>
</tr>
</tbody>
</table>

Rules have different reliability depending on their accuracy, formally, the ratio of the number of the forms correctly derived by the rule to the number of forms which meet the rule-environment. Reliability is adjusted so that reliability based on large data is better than reliability based on a few examples. The adjusted reliability is termed as “confidence”. Rules with high confidence values apply when speakers must synthesize new forms, for instance in the case of imperfect memory or loanword adaptation.

In the analysis of Korean noun paradigms, Albright (2005a) adopts the isolation form as the Base. Isolation form is not crucially more informative than other members of the noun paradigm, but it is the most frequently occurring one since case markers are usually omitted in natural speech in Korean. Under the idea of Albright (2005a), we may assume that the following multiple morphological rules are responsible for the occurrence of the stem-final obstruent alternations in Korean:

(31) Morphological rules for noun paradigms in Korean

<table>
<thead>
<tr>
<th>stem-final C</th>
<th>Rules</th>
<th>example pairs (isolation vs. topic form)</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. coronal</td>
<td>$t \rightarrow s / _+V$</td>
<td>ot os-in</td>
</tr>
<tr>
<td></td>
<td>$t \rightarrow t^h / _+V$</td>
<td>pat pat$^h$-in</td>
</tr>
<tr>
<td></td>
<td>$t \rightarrow c^h / _+V$</td>
<td>k’ot k’oc$^h$-in</td>
</tr>
<tr>
<td></td>
<td>$t \rightarrow c / _+V$</td>
<td>nat nač-in</td>
</tr>
<tr>
<td>b. labial</td>
<td>$p \rightarrow p / _+V$</td>
<td>ip ip-in</td>
</tr>
<tr>
<td></td>
<td>$p \rightarrow p^b / _+V$</td>
<td>ip$^b$-in</td>
</tr>
<tr>
<td>c. velar</td>
<td>$k \rightarrow k / _+V$</td>
<td>kuk kuk-in</td>
</tr>
<tr>
<td></td>
<td>$k \rightarrow k^b / _+V$</td>
<td>puæk puæk$^b$-in</td>
</tr>
</tbody>
</table>
As indicated by font size in (32), reliability (or confidence) of these rules differs depending on a difference in the type frequency of relevant words. As discussed above, the distribution of final obstruents of Korean noun stems shows the following hierarchies: in the case of coronal finals, s >> ch >> th >> c; and in the case of noncoronals, lenis >> aspirated. Based on these frequency hierarchies, the relative reliability of morphological rules are determined.

(32) Reliability differences of morphological rules for noun paradigms in Korean

a. t → s >> t → ch >> th >> c
b. p → p >> p

c. k → k >> k

The rules with high reliability/confidence may explain the creation of the innovative forms ending with [s] or noncoronal lenis stops.

In other analogy-based analyses (Ko 1989; Y. Kang 2003a,b), [s]-final forms are derived in a similar way based on isolation forms, which are adopted as the reanalyzed underlying form in Ko (1989) and the base of anticontrastive constraint in Y. Kang (2003b), as shown below:

(33) Other analogy-based approaches

a. Morphological rule: lt → [s] / __+V (noun paradigm) (trivially adapted from Ko (1989))

b. Anticontrastive constraint: t/\noun# → s/\nounV (t → s)

“If a noun ends in [t] in isolation form, change it to [s] before a vowel initial suffix.” (Y. Kang 2003b)

The reason why the isolation forms are the underlying form or base is their high token frequency. Ko and Y. Kang focus on the explanation of [s]-final forms, and, as far as I can see, it is very difficult to extend their analyses to variants ending in other obstruents.

These analogy-based analyses employ source-oriented rules/constraints, whose typical format is A → B. The source (A) is the isolation form with high token frequency, in which stem-final obstruents are neutralized, exhibiting no variation. The product (B) is a prevocalic form in which stem-final obstruents are contrastive, exhibiting variation. Let us now consider how to extend this type of analysis to variation patterns of the topicalized bound stems. The relevant facts are summarized below:
(34) Frequency and stem-final variation of the bound stems

<table>
<thead>
<tr>
<th></th>
<th>Isolation form</th>
<th>before –ha(ta)</th>
<th>before topic marker -in</th>
</tr>
</thead>
<tbody>
<tr>
<td>token</td>
<td>(almost) zero</td>
<td>high</td>
<td>very low</td>
</tr>
<tr>
<td>frequency</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>variation</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Coronal</td>
<td>lenis [tʰ]</td>
<td>aspirated [tʰ(ata)]</td>
<td>[s<del>cʰ</del>tʰ<del>c</del>t]</td>
</tr>
<tr>
<td>Labial</td>
<td>lenis [pʰ]</td>
<td>aspirated [pʰ(ata)]</td>
<td>[p~pʰ]</td>
</tr>
<tr>
<td>Dorsal</td>
<td>lenis [kʰ]</td>
<td>aspirated [kʰ(ata)]</td>
<td>[k~kʰ]</td>
</tr>
</tbody>
</table>

Unlike in the noun paradigm, bound stems, in principle, do not occur as isolation forms. Even if there might be a situation in which some bound stems are used in isolation, its use must be very limited. Even in such rare cases, final obstruents would appear as lenis stops due to coda neutralization, displaying no free variation. Bound stems are normally combined with the verb ha, and stem-final obstruents always appear as aspirated stops, due to the productive aspiration process where sequences of an obstruent followed or preceded by [h] become an aspirated stop. Only bound stems suffixed with the topic marker -in may show free variation of final obstruents.

Within the above mentioned analogy-based approaches, either isolation forms or pre-ha forms must be the source of the rule/constraint. There is no other possible option. If an isolation form is adopted as the source under the assumption that bound stems belong to nominal categories in the application of morphological and phonological rules/constraints, the same morphological rules/constraints would then apply to both nouns and bound stems, capturing their similarity in variation. However, there are serious problems with this option. First, as briefly discussed in section 3, it would be implausible to classify bound stems under consideration as nouns. Recall that many of them are mimetic words which should be classified as adverbs. In addition, all the previous analogy-based approaches take isolation forms of nouns as the base mainly because of their high token frequency. But, isolation forms of bound stems have basically zero token frequency, and thus they are not eligible for the base.

The other option, in which the compound verb form is the source, is also subject to serious problems. In this option, topicalized forms are determined by means of multiple morphological rules/constraints which crucially refer to stem-final aspirated stops as the rule source. Then the following rules may be proposed for the stems ending with coronal obstruents:

(35) Rules deriving topicalized forms of bound stems ending in coronal obstruents

(i) \( t^h \rightarrow s\text{in}\ h / \_\_\_\_\_\_\_ \text{ata} \)
(ii) \( t^h \rightarrow t^h \text{in}\ h / \_\_\_\_\_\_\_ \text{ata} \)
(iii) \( t^h \rightarrow c^h \text{in}\ h / \_\_\_\_\_\_\_ \text{ata} \)
(iv) \( t^h \rightarrow c \text{in}\ h / \_\_\_\_\_\_\_ \text{ata} \)
(v) \( t^h \rightarrow t \text{in}\ h / \_\_\_\_\_\_\_ \text{ata} \)
The question is how to explain similarity in stem-final obstruent variations between nouns and bound stems. To explain variations among noun forms ending with coronal obstruents, the following rules/constraints must be posited.

(36) Rules deriving topicalized forms of noun stems ending in coronal obstruents.

(i) \( t \rightarrow s \) in / __ #
(ii) \( t \rightarrow t^h \) in / __ #
(iii) \( t \rightarrow c^h \) in / __ #
(iv) \( t \rightarrow \ddot{c} \) in / __ #
(v) \( t \rightarrow t \) in / __ #

Notice that the two sets of rules/constraints in (35) and (36) are clearly distinct, especially taking different input forms, \([t^h]\) vs. \([t]\). Nonetheless, the outputs of their application would have to produce very similar patterns, i.e., similar relative preference/acceptability of stem-final variants. As discussed above, the relative productivity or reliability of morphological rules in (36) depend on the distribution of final obstruents of the noun stems. Then, it is not clear why the relative productivity or reliability of the rules for bound stems in (35) also depends on the distribution of a totally different morphological category, i.e., the noun. Consequently, there is no plausible way to extend the previous analogy-based accounts to the findings of the present study.

Finally, I want to emphasize that “source-orientedness” is not the only problem of the previous approaches. As mentioned above, all previous approaches are mainly concerned with s-final variant forms, and most of them can hardly be extended, without drastic revision, to deal with all the facts discussed in the present study, including the coexistence of five different coronal variants and their relative preference and acceptability. It seems that Albright’s model is better than other approaches in dealing with the expanded set of data, mainly because multiple rules may be posited and their reliability is determined by the relevant type frequency.

6. Conclusions
The conclusions of the present study are summarized as follows.

First, based on some recent survey studies, I have shown that patterns involved in free variation of stem-final obstruents in Korean nouns are more complicated than previously thought. Crucially, several different coronal obstruents are in variation with one another, and their usage preferences are different, mainly depending on the suffix.

Second, I have discussed frequency facts of Korean nouns to show that the relative preference among variants of the noun stems is matched by the type frequency of suffixed nouns. The same variation patterns and frequency matching have been observed in the results of a judgment survey employing a certain set of Korean bound stems. I have argued that all these findings and observations support the hypothesis
about the internalization of the variable lexical pattern.

Third, I have discussed the findings of the present study and their theoretical implications. Based on the interpretations of the findings, I have proposed a possible formal analysis of the variation patterns by adopting, with some modification, Optimality Theory. Language-specific, not universal, constraints are proposed to explain lexical frequency effects. And, variable, not fixed, rankings are adopted to explain free variation.

Finally, I have provided an in-depth discussion of how previous approaches to stem-final variation in Korean can deal with the observed patterns of the present study. I have shown that it would be very difficult to explain the variation patterns of Korean nouns and bound stems within the previous approaches, mainly because most of them focus on a limited set of data and they are source-oriented.
REFERENCES


Albright, Adam (2005a) Explaining universal tendencies and language particulars in analogical change. Ms, MIT.


Zuraw, Kie (2005) The Role of Phonetic Knowledge in Phonological Patterning: Corpus and Survey Evidence from Tagalog Infixation. MS, UCLA.
APPENDIX

Experimental sentences (Bound stems)\(^5\)

1. p-final stems

\begin{itemize}
  \item \textbf{pokčap} \textit{crowded}:
    \begin{itemize}
      \item \textit{Pusan-Nom a-little crowded-Top do look-Past-but clean-do-Con place-be-Past-SE}
        \textit{Pusan looked a little crowded, but it was a clean area.}
    \end{itemize}
  \item \textbf{takip} \textit{imminent}:
    \begin{itemize}
      \item \textit{in-one’s-own-way imminent-Top do-but, though anything buy-Part thing-Top none existent-SE}
        \textit{‘It’s urgent for me, but I shouldn’t buy anything without consideration.’}
    \end{itemize}
  \item \textbf{saŋkip} \textit{impatient}:
    \begin{itemize}
      \item \textit{ki salam saŋk’jak-i saŋkip-in ha-nte, kileto c’akhe}
        \textit{‘Although his personality is impatient, he is still nice.’}
    \end{itemize}
  \item \textbf{pikap} \textit{cowardly}:
    \begin{itemize}
      \item \textit{the person cowardly-Top do look-but though self work-Top well know-Con do}
        \textit{‘He looked cowardly, but he can do his own work well.’}
    \end{itemize}
  \item \textbf{čokip} \textit{rushed}:
    \begin{itemize}
      \item \textit{performance-Top ‘sufficiently-good’-do-Con rushed-Top do-Part need nonexistent-SE}
        \textit{‘Since his performance is sufficiently good, he doesn’t have to feel rushed.’}
    \end{itemize}
  \item \textbf{mihip} \textit{insufficient}:
    \begin{itemize}
      \item \textit{report-Nom yet insufficient-Top do-but, submit-Part thing-Top existent-Fut-SE}
        \textit{‘The report is still insufficient, but you may submit it.’}
    \end{itemize}
  \item \textbf{taptap} \textit{stuffy}:
    \begin{itemize}
      \item \textit{first hospitalization-thus room life-too stuffy-in do-but sleep-in well come}
        \textit{‘Since this is my first hospitalization, I feel stuffy, but I sleep well.’}
    \end{itemize}
  \item \textbf{sapsap} \textit{sorry}:
    \begin{itemize}
      \item \textit{fresh-sorry-do-Inter at-least sorry-Top do-SE}
        \textit{‘Are you asking whether I feel satisfied and, at the same time, sorry for it? At least I feel sorry.’}
    \end{itemize}
  \item \textbf{k’amucapcap} \textit{‘a little dark’}:
    \begin{itemize}
      \item \textit{face-Nom ‘a-little-dark’-Top do-but pretty-SE}
        \textit{‘The face is a little dark but it’s pretty.’}
    \end{itemize}
  \item \textbf{kapkap} \textit{‘stuffy’}:
    \begin{itemize}
      \item \textit{underwear wear-because stuffy-Top do-but not-wear-Part thing-than warm do-SE}
        \textit{‘Since I wore underwear, I feel stuffy, but it’s warmer than not wearing it.’}
    \end{itemize}
  \item \textbf{č’ipč’ip} \textit{‘unsatisfied’}:
    \begin{itemize}
      \item \textit{dyeing do-Fut-thus unsatisfied-Top do-but do-Con-Top wish-and almost two-year-Acc patient-Past-SE}
        \textit{‘I don’t feel comfortable about dyeing, but I want to dye my hair. I have been waiting for two years.’}
    \end{itemize}
\end{itemize}

2. s-final stems

\begin{itemize}
  \item \textbf{pisis} \textit{similar}:
    \begin{itemize}
      \item \textit{wančənhi katb’ci-nin anh-talato pisis-in he-jači}
        \textit{completely same-Top not-though similar-Top do-shoud}
      \item \textit{‘Even if (they) are not completely identical, they should be similar.’}
    \end{itemize}
  \item \textbf{k’ek’is}:
    \begin{itemize}
      \item \textit{kileto k’ek’is-in he poi-čanha}
    \end{itemize}
\end{itemize}

\(^5\) (Top = Topic marker, Nom = Nominative case marker, Acc = Accusative case marker, Con = Converb, Part = Participle, SE = Sentence Ender, Past = Past tense suffix, Fut = Future tense suffix, Inter = Interrogative suffix)
3. k-final stems

- **'clean'**
  - “It looks clean, though.”
  - “Clean Top do look-SE

- **'t'at'is'**
  - This clothing shape-Top pretty not-but warm-Top do
  - “This clothing looks ugly but it keeps me warm.”

- **'warm'**
  - “We lost the game, but we are open and aboveboard nonetheless.”
  - “This clothing shape-Top pretty not-but warm-Top do
  - ‘The lady is generous.’

- **'aboveboard'**
  - “The illness is serious, but it’s treatable.’
  - ‘The injured part is hardened right now, but it will be cured soon.’

- **'stimulating'**
  - ‘The boy friend is honest, but he is weak-minded.’

- **'dignified'**
  - “He is young, but he behaves in a dignified manner. ’

- **'satisfied'**
  - “The personality is relaxed, though.’

- **'pleased'**
  - “Since I’ve finished the work, I feel satisfied.’

- **'relaxed'**
  - “This clothing looks ugly but it keeps me warm.’

- **'concern-SE’**
  - “The movie stimulating-Top do-but level-Top nonexistent-SE

- **'pitiful'**
  - “The movie is stimulating, but it’s low quality.’

- **'love-Top'**
  - “The love between the man and the lady in the movie is pitiful.’

- **'warm'**
  - “The person is decent in manner.’

- **'decency-Top'**
  - “Since I’ve finished the work, I feel satisfied.’

- **'empathy-SE'**
  - “The boy friend is honest, but he is weak-minded.’

- **'express-DO'**
  - “The name is pretty, and she is clever, but she doesn’t know how to express her emotion.’

- **'aboveboard'**
  - “The person is decent in manner.’

- **'decency-Top'**
  - “Since I’ve finished the work, I feel satisfied.’

- **'empathy-SE'**
  - “The boy friend is honest, but he is weak-minded.’

- **'express-DO'**
  - “The name is pretty, and she is clever, but she doesn’t know how to express her emotion.’

- **'belowboard'**
  - “The lady is generous.’

- **'concern-SE’**
  - “The movie stimulating-Top do-but level-Top nonexistent-SE

- **'pitiful'**
  - “The movie is stimulating, but it’s low quality.’

- **'love-Top'**
  - “The love between the man and the lady in the movie is pitiful.’

- **'warm'**
  - “The person is decent in manner.’

- **'decency-Top'**
  - “Since I’ve finished the work, I feel satisfied.’

- **'empathy-SE'**
  - “The boy friend is honest, but he is weak-minded.’

- **'express-DO'**
  - “The name is pretty, and she is clever, but she doesn’t know how to express her emotion.’

- **'belowboard'**
  - “The lady is generous.’
‘Although the movie was horrible, I still want to see it again.’

‘Recently, the weather change is drastic, but it’s generally warm.’

‘Since aunts are coming over, it will be noisy.’

‘My wardrobe is full of clothes, but I would like to wear none of them.’

‘He looks lively, but he is still sick.’