http://dx.doi.org/10.17959/sppm.2016.22.3.559

Phonological optimization through entropy minimization*

Tae-Jin Yoon
(Sungshin Women's University)

Yoon, Tae-Jin. 2016. Phonological optimization through entropy minimization. *Studies in Phonetics, Phonology and Morphology* 22.3. 559-581. This study aims at accounting for suppletive allomorph selection in the Korean noun particle system using an Information Theoretic approach. Approaches that utilize phonological optimization fail to account for suppletive allomorph selection in Korean noun particles. It is likely that the phonological optimization seen in a subset of the noun particle system (i.e., Nominative, Topic, and Accusative, among others) may be a by-product that has arisen in the minimization process of entropy or uncertainty in the Korean nominative particle system. The approach which is message-based draws an implication that optimal sound patterns do not exist for their own sake, but are the result of optimizing communication. **(Sungshin Women's University)**

Keywords: stem selection, allomorph selection, entropy, phonological optimization, minimization of entropy

1. Introduction

This study aims at accounting for allomorph selection in the Korean noun particle system using Information-theoretic approach, with a particular focus on conditional entropy. Allomorph selection in the noun particle system in Korean appears to be phonologically conditioned. For example, some suffixes in Korean have a sonorant-initial allomorph that occurs after vowel-final stems and a vowel-initial allomorph that occurs after consonant-final stems (Ahn 1996, Sung 2005, G. Kim 2006, Suh 2006, Lee 2008, among others). A nominative allomorph *–ka* occurs after stems ending with a vowel such as *kho* 'nose' and the other allomorph *–i* happens after stems ending with a consonant such as *mom* 'body. Seemingly, the choice of the allomorphs appears to be optimized phonologically by making the syllable structure

* I thank three anonymous reviewers for their constructive comments and suggestions. The paper improved much with the help of the reviewers. The work was supported by the Sungshin University Research Grant of 2016, to which I am grateful.

an unmarked CV structure (Sung 2005, Paster 2006, Suh 2006, Lee 2009, Nevins 2010).

Due to the intricate interplay between morphology and phonology, the seemingly phonologically-conditioned suppletive allomorphy has been accounted for in a number of different ways rather than solely by phonological rules or constraints (Paster 2006, Bye 2007). For example, Sung (2005) categorized the Korean noun particle and verb conjugation systems, as in (1), into three types: phonologically conditioned allomorphs, allomorphs selected by *l*-final stems, and allomorphs of the conjunctive suffix -*wa*/-*kwa*. She maintained that even though phonological environments could explain the first type of, the other types could not be accounted for without referring to the interaction between morphology and phonology.

(1) Korean allomorphy alternation in noun particle[1] (Lee 2008: 68)

noun particle	**allomorphy**
Nominative	i/ka
Topic	ɨn/nɨn
Accusative	ɨl/lɨl
Conjunctive	wa/kwa
Goal/Instr.	lo/ɨlo

Previous analyses include rule-based approaches (Odden 1993, Sung 2005), phonology-based OT approaches (Sung 2005, Suh 2006, Lee 2008), and OT approaches with an emphasis on the interplay of phonology and morphology (G. Kim 2006).

In this paper, I propose an alternative approach to the suppletive allomorph selection. In particular, I propose that entropy can be used as a criterion of selecting the correct allomorph and that it can be incorporated as an allomorphic selection mechanism. Entropy is derived from the Information Theory (Shannon 1948) and is used to quantify predictability or uncertainty. When used in grammatical relations, entropy can provide a way of formally describing and comparing even those

[1] The discussion will be limited to the noun particle system presented in Table 1, as these are the most widely discussed cases in the literature on stem selection in Korean. However, the approach presented in this paper can be extended to other allomorph cases, even though the result may not be as clear as the one that will be reported later in this paper.

grammatical relations which have been treated as problematic for traditional descriptions of grammatical components including allomorphic relations (Hall 2009).

The paper is organized as follows: Section 2 presents some cases of phonologically conditioned allomorphs, some of which are not phonologically optimized. Section 3 introduces Entropy and Conditional Entropy, and then applies the concepts to the Korean nominative system. Section 4 concludes the paper.

2. Phonologically-Conditioned Allomorph Selection

In this section, cases of phonologically-conditioned allomorph selection will be examined, starting from the cases of phonological optimization to the cases of non-phonological optimization.

Lots of allomorph-selection cases are phonologically conditioned. Kager (1996) divides phonologically conditioned allomorphs into two classes: fully-conditioned and partially-conditioned allomorphs. In the case of the fully-conditioned allomorphs, complementary distribution follows completely from phonological principles such as the avoidance of ill-formed syllables, without any need for a supplementary morphological statement. In the case of partially-conditioned allomorphs, phonology accounts for the distribution of only one allomorph, while the distribution of the other allomorph must be due to morphological principles such as a preference for one allomorph over another.

A distinction between bases ending in a consonant and those ending in a vowel seems to be a good example of phonologically conditioned allomorph selection. Several languages including Northwestern Catalan, Morocan Arabic, and Armenian are of this type, and Mascaró (2007) goes further to state that most of the cases reported as external allomophy are related to syllable structure.

Definite articles in Northwestern Catalan are sensitive to the type of segments that follows the articles. The Northwestern Catalan masculine singular definite articles are either *lo* or *l*: *l* appears before vowels and *lo* appears before consonants (Mascaró 2007). Allomorphic choice in Catalan is phonologically natural since it determines a less-marked CV syllable structure.

(2) Northwestern Catalan

l ámo	'the owner'	**lo** pá	'the bread'
l uɾaká	'the hurricane'	**lo** míte	'the myth'

In Moroccan Arabic, the 3rd sg. masculine pronominal clitic is either [h] or [u]. [h] appears after vowels, and [u] appears after consonants (Mascaró 2007)[2].

(3)	xtˤa	'error'	xtˤa-h	'his error'
	ʃafu	'they saw'	ʃafu-h	'they saw him'
	ktab	'book'	ktab-u	'his book'
	ʃaf	'he saw'	ʃaf-u	'he saw him'

Another example of this phenomenon is found in Armenian (Vaux 1998), where the definite article is expressed as a -*ə* suffix when the stem is consonant-final, but as -*n* when the stem is vowel-final.

As a case of partially-conditioned allomorphy, Kager (1996) used the genitive marker system in Djabugay (Patz 1991) as an example.

(4) (a) /-n/ after bases ending in a vowel
guludu 'dove' guludu-n 'dove-GEN'
(b) /-ŋun/ after bases ending in a consonant
gaɲal 'goanna' gaɲal-ŋun 'goanna-GEN'

The choice of the proper allomorph in (4) is not fully but partially predictable from the syllable structure. That is, although the genitive allomorph /-n/ may be accounted for by syllabification, syllabification alone is insufficient to explain why the allomorph /-ŋun/ cannot be attached to vowel-final bases. Even if /ŋun/ is attached to the bases ending in a vowel (e.g. ?*guludu-ŋuni*, it does not violate any syllable well-formedness.

There are cases of allomorph selection which are irregular and internally conditioned in the lexicon. The perfective suffix -oh/-ɛh in Tzeltal fits the case. In Tzeltal, -oh appears after monosyllabic stems and -ɛh appears after polysyllabic stems (Mascaró 2007).

(5) Tzeltal perfective suffix –oh/-ɛh
j-il-**oh** 'he has seen something'

[2] Mascaró (2007) refers to this kind of allomorph relation as 'external allomorphy. In external allomorphy, the choice of allomorphs contributes to optimize the surface forms. The converse case is 'internal allomorphy,' in which case the choice of allomorphs has no optimizing character at all (Lee 2009).

s-nu͡ts-**oh** 'he has chased something'
s-mak'lij-**ɛh** 'he has listened to something'
s-tikun-**ɛh** 'he has sent something'

Allomorphic choice in Tzeltal is arbitrary and does not improve the resulting structure regarding unmarkedness of syllable structure.

The Haitian Creole allomorphy is considered to be "perverse" from the perspective of syllable-structure markedness. Affixing *–la* to consonantal-final hosts creates syllable codas and affixing *–a* to vowel final hosts creates hiatus between syllables. Both of these results are non-optimal, and both of these problems would disappear if the reverse distribution of allomorphs obtained (Embick 2009: 5)[3].

(6) Haitian Creole definite suffix –la/-a
liv-**la** 'book'
tu-**a** 'hole'

Since we have examined some cases in which the allomorph selection is not always phonological optimized, we will review a few approaches that attempt to deal with the suppletive allomorph selection.

Now the question to address is at what stage in the phonological derivation allomorph selection takes place and what is the mechanism of allomorph selection? The existence of phonological conditioning suggests that grammatical mechanisms in the phonological component are involved in choosing a proper allomorph rather than listing of allomorphs in the lexicon. Therefore, given this premise, phonological relations need to be properly redefined for allomorph selection to be successful.

A pure phonological approach to the phenomenon of phonology-morphology interface is to posit deletion and insertion rules with one of the allomorph set to be an underlying form. For example, Odden (1993: 134): in the choice of the nominative markers -*ka* and -*i*, one may posit -*ka* as an underlying form for the nominative marker, and try to derive -*i* from -*ka* by two deletion rules and one insertion rule[4].

---

3 There is a possibility that the apparent non-optimizing nature on syllable structure may be reinterpreted to be optimizing in the alignment or syllable contact markedness (Suh 2006).

4 Phonologically conditioned suppletive allomorphs cannot be traced back to a single underlying form. Motivated phonological relation cannot be found between allomorphs. But their distribution is strictly governed by phonology (Lee 2009).

However, this type of approach cannot account for the naturalness of the rules involved, due to the lack of surface resemblance between two allomorphs, -*i* and -*ka*.

Kager (1996) propose a solution to the partially-conditioned allomorph selection in the framework of OT. In this approach, phonological effects in morphology are modeled by ranking phonological (P) over morphological (M) constraints (i.e. 'P >> M' in the sense of McCarthy and Prince's (1993)[5]. That is, a phonological constraint outranks a morphological constraint in that a phonological constraint *COMPLEXCODA ("No Complex codas are allowed") is ranked higher than a morphological constraint GENITIVE=/-n/ ("The genitive is marked by /-n/)"). The prediction that can be made according to this approach is that the selected allomorphs should serve to "optimize words with respect to phonological constraints that are independently motivated elsewhere in Universal Grammar (Paster 2006)."

Since many suppletive allomorphs do not have any phonological bearings, Lapointe (1999) proposed Multiple Input Hypothesis (MIH). Because it is reasonable to assume that the suppletive allomorphs cannot be reduced to single input, MIH allows all the suppletive allomorphs in the input and leaves the final choice to the interaction of constraints. As for choosing the correct allomorph, Paster (2006) maintains that the existence of non-optimizing suppletive allomorphy demonstrates the need for a mechanism other than output optimization to handle PCAS (Phonologically Conditioned Allomorph Selection). A *subcategorization* approach captures the fact that there are cases of PSCA where phonological well-formedness considerations have no bearing on the choice among allomorphs in a given environment. The main idea is that the representation of an affix includes requirements for stems to which it will attach.

Paster's (2006, 2008) lexical subcategorization approach relies on the lexical information that stipulates the distribution of internal allomorphs. She does not discuss how the subcategorization information is implemented into constraints. There might be two possibilities: one is to have a constraint that penalizes the allomorph outside of the defined subcategorizational information, another is to translate the lexical subcategorization information into allomorph-specific constraints.

---

5 This is not the only solution. For previous proposals are neatly summarized in Lee (2009) and Lee (2010). Previous approaches include (i) Process Priority (McCarthy and Wolf 2005), (ii) Morpheme as constraint (Kager 1996), (iii) Morpheme markedness (Łubovicz 2006), and (iv) Shape Priority (Mascaró 2007).

G. Kim (2006) takes on the subcategorization approach to the allomorph selection with a flavor of Lexical Phonology and the framework of subcategorization. Her proposal is that “each allomorph of alternation suffixes is subcategorized with respect to the features of its preceding stem-final segment (G. Kim 2006: 265).” The inconsistent behavior of /l/ is explained with the feature specification of /l/ as [+vocalic]. She illustrates the process of selecting allomorphs as follows, taking *sal-* ‘to live’ and Effective for example.

(7) Step 1: stem-final segment identification
stem: sal-, /l/: [+vocalic, +consonantal]
Step 2: Suffix identification and allomorph selection
Effective → [+/- vocalic]; -ni [+vocalic] + ____
-ɨni [-vocalic] + _____
Step 3: (i) rule-based analysis: derivation
(ii) OT-based analysis: candidate evaluation
(G. Kim 2006: 280)

Lee (2009) proposes a new constraint Default which states that “a phonologically simpler allomorph is preferred.” According to Lee (2009), an allomorph is phonologically simpler, if it has less number of segments and/or if it is less marked regarding feature composition. The default form or the preferred form is selected by the phonological simplicity measures encoded in DEFAULT. There is a problem with DEFAULT, however. Boyd 2006, in his analysis of Italian article *il/lo* alternation, has to postulate a constraint, *lo (i.e., do not use /lo/), to capture the distributional generalization. The postulated constraint implies that /*il*/ is the preferred constraint. Thus, Lee (2009: 425) needs to make /*il*/ as the default allomorph over /*lo*/. He asserts that “/*i*/ is less marked than /*o*/, hence /*il*/ is less marked than /*lo*/,” without taking into consideration the fact that /lo/ (which is CV) is less marked than /*il*/ (VC) in terms of syllable structure.

A problem in the OT-based approaches to the suppletive allomorph selection is that there are no consistent constraint rankings that account for all Korean nominative particle system. The OT-based account has to stipulate constraints that are not motivated elsewhere in the grammar. In the next section, an alternative approach to the suppletive allomorph selection will be presented using an Information Theoretic approach.

3. Entropy Minimization

An alternative to the previous approaches to phonologically conditioned allomorph selection comes from the assumption that language is a system of information transmission and adopt a message-based approach rather than a sound-based approach (Hume et al. 2016). Hume et al. (2016) take the position that the message as a unit of meaning in spoken language is transmitted by an acoustic/auditory speech signal. The communication system can be approximated by "taking into account of uncertainty associated with selecting the outcome of a transmitted message (Shannon 1948)." Uncertainty can be quantified using the concept of Entropy in Information Theory (Shannon 1948), and can be used to calculate predictability of distribution of phonological units, as in Hall (2009).

Entropy is defined as the weighted average of the uncertainty for all possible outcomes (x's) of a random variable (*X*) (Shannon 1948, Pierce 1980, Malouf 2009, Hong 2014). Entropy can be quantified using the formula in (8):

(8) Definition of Entropy

$$H(X) = -\sum_{x \in X} p(x) \log_2 p(x)$$

Note that *p(x)* is the probability that the outcome of *X* is *x*.

Figure (1) illustrates the value of entropy in the case of two outcomes of a variable:

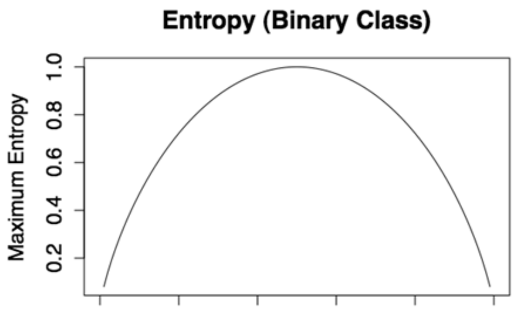

Figure 1: Entropy in the case of two possibilities with probabilities *p* and (1-p) (Shannon 1948: 50). When the two class has the same probability, i.e., *p* = 0.5, the entropy value reaches its maximum value, or the maximum uncertainty.

As shown in Figure (1), if outcomes are binary, the maximum entropy is 1, and this maximum entropy represents the state of affairs that is the most uncertain. This state of maximum uncertainty is made when the two classes are the same probability.

In phonologically conditioned allomorphs, predictability of distribution is crucial (Hall 2009). There are instances, however, where the phonological predictability is not clear-cut. Due to the existence of vague instances, one either argues for the position that allomorphs are listed in the lexicon or asks for a formal mechanism for allomorph selection that is robust enough to be employed in uncertain environments.

As a simple example illustrating the function of entropy, let's consider an example from the Korean nominative particle. The nominative suffix in Korean shows an allomorphic alternation between -*i* and -*ka*, as in Table 1 (Lee 2008). In this paper, the Korean examples are presented with the Yale Romanization system.

Table 1: Nominative particle allomorphs

Stem	*i*-particle	*ka*-particle	Gloss
wang	wang-i	*wang-ka	'king-Nom'
mul	mul-i	*mul-ka	'water-Nom'
so	*so-i	so-ka	'cow-Nom'
pi	*pi-i	pi-ka	'rain-Nom'

Let's assume, at the moment, that we don't know anything about the nominative particle except that it has two allomorphs -*i* and -*ka*, and that the probabilities of all possible allomorphs of the nominative particle sum to 1. That is, in the case of X being the nominative particle:

(9) p(Nom) = p(Nom=i)+p(Nom=ka)=1

The entropy H for the Nominative particle is:

(10) $H(Nom) = -(p(N{=}i) * \log_2 p(N{=}i) + p(N{=}ka) * \log_2 p(N{=}ka))$

Given the data in Table 1, the entropy of the Korean Nominative particle is:

(11) $H(Nom) = -(0.5 * \log_2(1/2) + 0.5 * \log_2(1/2)) = 1$ (bit)

The value of entropy can be interpreted in a couple of ways (Malouf 2009): It can be (1) the average number of bits required to store the value of X, or (2) the average number of yes-or-no questions that one has to ask to guess the correct value of X. In the case of the Korean Nominative particle, we need 1 bit to store the correct allomorph of the Nominative marker, or we need to ask at least once to guess correctly whether the allomorph of the Nominative marker is -*i* or –*ka*, which is intuitively correct. If we need to ask at least once to find out what allomorph the Nominative particle bears, it means the selection of the Nominative allomorph, when nothing is known, is very unpredictable or uncertain. This intuitive interpretation can be viewed through the graphical representation of entropy when a variable has only a binary choice, as in Figure 1.

However, it is rare to find events that have the same probability. For example, Table 2 shows the distribution of the nominative particle obtained from one file in Sejong21 Corpus developed through the 21st Sejong Project[6].

Table 2: The distribution of -*i* and -*ka* obtained from a file in Sejong Corpus

Nom. suffix	Frequency	p(x)	$p(x)*\log_2*p(x)$
-*ka*	153/432	0.355	-0.53
-*i*	279/432	0.645	-0.40

The frequency of -*ka* calculated in the file is much higher than that of -*i*. As shown in (5), the probabilities between -*ka* and -*i* are asymmetric, and the value of entropy is lowered to 0.93 decreasing the uncertainty.

(12) $$\begin{aligned} H(Nom) &= -(0.35*\log_2(0.35) + 0.64*\log_2(0.64)) \\ &= -(-0.53-0.40) \\ &= 0.93 \text{ bit} \end{aligned}$$

6 Only a single file, and hence a small number of tokens, is used in the corpus, because the purpose of the demonstration is to show how the entropy can be calculated from a corpus, rather than examine the whole corpus regarding the distribution of –*ka* and –*i*. Regardless of the distribution of –*ka* and –*i*, the nature of complementary distribution of these two allomorphs would make the argument in the paper valid. The number of tokens extracted from the whole Sejong21 corpus is presented in Table 3.

This illustrates that probability affects the uncertainty inherent in the Nominative particle.

Table 3 presents the total number of tokens is extracted from the morphologically tagged portion of the Sejong21 corpus (H. Kim 2006). The morphological tags such as JX (for Topic), JKS (for nominative particle), JKO (for Accusative), JKB (for adverbial particles including Goal/Instrument), and JX (for Conjunctive) are identified first, and then each allomorph is computed from the identified tags.

Table 3: Allomorph alternation of noun particles in Korean. The number of tokens is calculated from the written corpus in Sejong21 Corpus.

Particle	allomorph	Tokens #	allomorph	Tokens #
Nominative	-i	350,759	-ka	224,939
Topic	-ɨn	253,248	-nɨn	317,295
Accusative	-ɨl	503,312	-lɨl	287,995
Goal/Instrument	-ɨlo	150,498	-lo	110,040
Conjunctive	-wa	46,541	-kwa	70,481

In the table, the nominative, topic and accusative markers have two allomorphs, one with an initial vowel and the other with an initial consonant. Also in the table, the number of tokens for each allomorph is presented, which may be needed to calculated entropy. The calculated entropy is 0.96 for nominative, 0.99 for topic, 0.94 for accusative, 0.98 for goal/instrument, and 0.96 for conjunctive. Not surprisingly, all of the values, which is very close to 1 (*i.e.*, maximum uncertainty) suggests that the binary allomorphs are in near complementary distribution.

It is not always the case that entropy is used only for binary cases. There are cases where the number of classes exceeds 2[7]. For example, the allomorphs for the regular past tense are chosen out of *-d*, *-t*, and *-əd*. The morphological passive morphemes in Korean is one of the four *-i*, *-hi*, *-li*, and *-ki*. Figure 2 illustrates the relationship

[7] A reviewer suggests a three-way contrast, rather than a binary opposition. For example, nominative marker can be omitted in Korean, thus a null allomorph can form an opposition together with *–i* and *–ka.* This is a valid point. But in a corpus study like the current paper, it is quite difficult and time-consuming to decide where a potential null allomorph occurs and what type of particle (for example, whether nominative or accusative) the function of the omitted allomorph may be. Thus, the paper restricts its scope to the choice of allomorph to binary cases only.

between the number of classes (on the x-axis) and the value of maximum entropy (on the y-axis in the right-hand side) and the probability (on the y-axis in the left-hand side). If the number of classes is *n*, then the probability is equal to $p=1/n$. This is illustrated on the y-axis on the left-hand side. As the number of classes increases, the value of maximum entropy also increases.

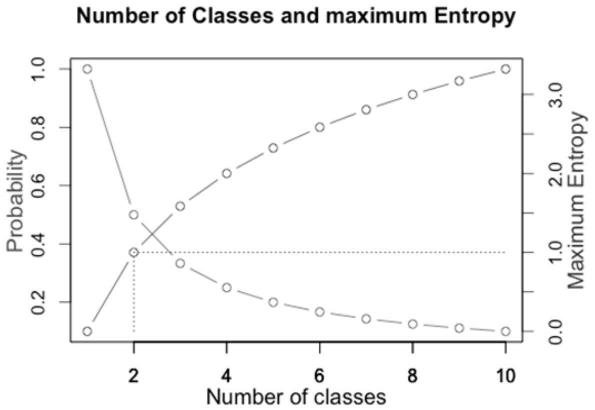

Figure 2. Entropy reaches maximum value when all classes have equal probability. The values of maximum entropy are plotted for different number of classes *n*, where probability is equal to p=1/n. In this case, maximum entropy is equal to $-n*p*\log_2 p$.

Entropy has been applied to a number of phonological problems, among others, in deciding the allophones and phonemes (Hall 2009), in deciding passive allomorphs in Korean (Yoon 2012), and in accounting for the vowel epenthesis in Korean loanword phonology (Lee 2015).

Hall (2009) noted a long-standing debate about the status of the vowels [aɪ] and [ʌɪ] in Canadian English. She asked the question whether they should be considered allophonic or contrastive. Because they are predictably distributed in most environments, they should be considered allophonic. However, because they are also unpredictably distributed before flap in near minimal pairs, they should be considered contrastive. The Canadian Raising is only one of many cases that are problematic for the usual classification of being allophonic or contrastive. Using PPRM (Probabilistic Phonological Relationship Model) for Canadian English, she demonstrates that the relationship between [aɪ] and [ʌɪ] is one of partial predictability.

Specifically, [aɪ] and [ʌɪ] have a relatively high uncertainty of approximately 0.95 in the pre-flap environment, but the impact of this environment is relatively small, so that the overall uncertainty of the pair is very low, near 0.01 (Hall 2009).

Yoon (2012) extended the unconditional entropy to the conditional entropy in accounting for the allomorph selection of the Korean Passive. Morphological passive construction is derived by concatenating a lexically specified verb and one of the four allomorph, -*i*-, -*hi*-, -*li*-, and –*ki*-. Only one of the four allomorphs is allowed, depending on the type of the verb. As noted in Ahn (1996), the stem-final *t* and *c* occur before both –*ki*- and –*hi*-, making the *t*-final and *c*-final stems exceptional cases, as in (13)

(13) Seemingly exceptional cases of the passive allomorphs –*ki*- and –*hi*-

-*ki*		-*hi*	
t'u**t**-ki-ta	'be picked'	ta**t**-hi-ta	'be closed'
c'o**c**-ki-ta	'be chased'	me**c**-hi-ta	'be tied'
c'i**c**-ki-ta	'be torn'		

Yoon (2012) implemented a computer simulation using the data from Ahn (1996). Ahn (1996) makes available a total of 148 morphological passive items. The list of passive items is compiled based on 3 Korean dictionaries. Each passive item was manually converted as a feature vector taking from onset, nucleus, coda of the final syllable in the passive verbal stem, along with the target attribute of the four types of passive allomorphs. Taking as the input the feature vectors of the 148 morphological passive items, the computer simulation computed conditional entropy for each verbal stem for allomorph. The conditional entropy values in the syllable coda of the verbal stem appeared to be informative since most codas have zero entropy value. It was also evident that uncertainty value occurred when the coda of the stem-final syllable is either –*t* or –*c*, as the two non-zero entropy values of 0.65 (in the case of –*t*) and 0.97 (in the case of –*c*) illustrated[8].

Lee (2015) applied the notion of conditional entropy to the problem of epenthesis in the context of word-ending affricate consonants. As noted in Oh (1992), Korean

8 A reviewer raised a question whether the allomorph can be correctly selected using conditional entropy for the passive allomorphs in Korean. Yoon (2012) demonstrated that the correct selection can be made by combining iterative computation of conditional entropy with decision tree.

specific dependency between preceding affricates and accompanying vowels, such that instead of typical [ɨ] epenthesis, high front vowel [i] emerges when affricate consonants occur word-finally (e.g., page [pʰeidʒi] vs. pet [pʰetʰɨ]. Lee (2015) accounts for the emergence of [ɨ] rather than [i] in certain loanwords as the result of conditional entropy.

So far, we have examined some previous studies of phonological problems that are approached to from the perspective of entropy. Despite the fact that entropy can be used to quantify uncertainty, it does not seem to function well enough to be used as a mechanism for allomorph selection.

Conditional entropy is, on the other hand, well suited for this purpose of selecting the correct allomorph. Conditional entropy can be viewed as a way of reducing the uncertainty is adding knowledge to the system. The conditional entropy of Y given X expresses the average additional information of Y, given that X is known, as formulated in (14) (Shannon 1948, Pierce 1980):

(14) Definition of Conditional entropy

$$H(Y|X) = -\sum_{x\in X}\sum_{y\in Y} p(x)p(y|x)\log_2 p(y|x)$$

In the previous examples of the Nominative particle, we assumed that we do not know anything about the nominative marker except that the Nominative particle has two allomorphs -*i* and -*ka*. Now let's introduce additional knowledge into the system. Specifically, let's assume that we know at least that -*i* occurs after a consonant-final stem, and -*ka* occurs after a vowel-final stem, as in Table 3.

Table 3: Nominative particle allomorphs: -*i* after a C-final stem, -*ka* after a V-final stem

Stem	*i*-particle	*ka*-particle	Stem-final	Gloss
wang	wang-i	*wang-ka	C	'king-Nom'
mwul	mwul-i	*mul-ka	C	'water-Nom'
so	*so-i	so-ka	V	'cow-Nom'
pi	*pi-i	pi-ka	V	'rain-Nom'

As in (15), it becomes clear that by adding knowledge to the system and hence by using conditional entropy, we can reduce the uncertainty. In this very simple case of the Korean nominative marker, by adding the knowledge of the status of the preceding segments, i.e., whether the preceding segment ends with a consonant or a vowel, maximum certainty is achieved.

(15) a. H(Nom | SF=V) = H(Nom=ka | SF=V) + H(Nom=i | SF=V) = 0
b. H(Nom | SF=C) = H(Nom=ka | SF=C) + H(Nom=i | SF=C) = 0
c. H(Nom | SF) = H(Nom | SF=V) + H(Nom | SF=C) = 0

In (15a), H(Nom=ka | SF=V), which means the conditional entropy in which Nominative is realized as -*ka* when stem final segment is a vowel, is 0 because there is no occurrence of –*ka* when the stem final segment is a vowel, and H(Nom=i | SF=V) is 0 because all the occurrence of the allomorph is –*i* when the stem final segment is a consonant. The reverse is true for (15b). Thus, when we add the two of (15a) and (15b), we get (15c), which states that then entropy value of nominative particle given the status of stem final segment is 0.

The conjunctive has apparently contradictory distributions compared to the distribution of other markers. For example, the conjunctive allomorph -*kwa* is selected when the stem-final segment ends with a consonant, and the allomorph -*wa* is selected when the stem-final segment ends with a vowel, as seen in (16).

(16) Korean conjunctive suffix allomorph selection (Nominative counterparts are in the parenthesis for comparison)

a.	mom-kwa	*mom-wa	‘body-CONJ’
	(mom-i	*mom-ka	'body-NOM’)
b.	kho-wa	*kho-kwa	‘nose-CONJ’
	(*kho-i	kho-ka	‘nose-NOM’)

Thus, the conjunctive particle behaves in quite the opposite way to other particles such as the nominative particle. This opposite behavior makes the explanation of the noun particles based on the phonological optimization problematic. It should be noted that even in the nominative particles –*i*/-*ka*, the choice of allomorph does not always result in optimizing effect. Lee (2008) notes that when stems ending in velar nasal (e.g., *wang* ‘king’) and the noun particle –*i* are concatenated, the resulting syllable structure is CVC.V, due to the phonotactic constraint that does not allow velar nasal in the syllable onset position. The lack of consistency in the noun particles makes some researchers entertain the possibility of treating allomorphic selection as morpho-lexical, rather than phonological, properties.

If we approach the allomorph selection from the perspective of conditional entropy, this seemingly contradictory observation does not pose any challenge.

Table 5: Allomorphic alternation of Conjunctive marker: /kwa/ after a C-final stem, /wa/ after a V-final stem

Stem	kwa-particle	wa-particle	Stem-final	Gloss
wang	wang-kwa	*wang-wa	C	‘king-CONJ’
mul	mul-kwa	*mul-wa	C	‘water-CONJ’
so	*so-kwa	so-wa	V	‘cow-CONJ’
pi	*pi-kwa	pi-wa	V	‘rain-CONJ’

As shown below in calculation of the conditional entropy (17) on the basis of the data in Table 5 for the conjunctive marker, the conditional entropy values give rise to 0, which indicates that there is no uncertainty in the selection of the proper allomorph. In (17), regardless the stem-final segments, the total entropy for each noun particle sums up to be zero, whose value indicates the lack of uncertainty in the choice of suppletive allomorphs.

(17) a. H(CONJ | SF=V) = H(CONJ=ka | SF=V)+H(CONJ=i | SF=V) = 0
b. H(CONJ | SF=C) = H(CONJ=ka | SF=C)+H(CONJ=i | SF=C) = 0
c. H(CONJ | SF) = H(CONJ | SF=V) + H(CONJ | SF=C) = 0

The other allomorphs in Table 3 result in the value of conditional entropy as 0. It is seemingly not appealing to say that conditional entropy can be used instead of previous rule- or constraint-based approaches, because such previous mechanisms can account for the selection of each nominative allomorph due to the nature of near complementary distribution of each of the allomorphs listed in Table 3.

The real problem in Korean noun particle system is that the system as a whole is very hard to be accounted for by the previous approaches without introducing *ad hoc* propositions. For example, the nominative suffix -*ka* occurs after stems such as *kho* ‘nose’, and -*i* occurs after stems such as mom ‘body.’ The choice of the allomorphs, thus, seems to optimize phonologically by making the syllable structure as the unmarked CV structure (Sung 2005, Suh 2006, Lee 2008, Nevins 2010). As is the case with *kho-ka* being syllabified as *kho.ka* (CV.CV) and *mom-i* as *mo.mi* (CV.CV). If different allomorph is chosen, this unmarked syllable structure cannot be maintained. For example, improper selection of the allomorph in the case of **kho-i* results in CV.V and in the case of **mom.ka* as CVC.CV, both of which contain the marked V and CVC syllable structure, respectively. As is illustrated using the nominative particle, conditional entropy can be utilized to select an allomorph. Thus, evaluation of two approaches is made by looking at the consistency of treatment in the account of the other noun particles.

The topic and accusative particles show the same pattern as the nominative particle, but goal/instrument particle *-lo*/*-ɨlo* and conjunctive particle -<u>*wa*</u> and -*kwa* pose a challenge to the generalization made based on syllable structure (Sung 2005, Suh 2006, Lee 2008, Nevins 2010). Goal/instrumental particle *-lo/-ɨlo* shares resemblance in its shape to accusative particle *-ɨl*/*-lɨl.* Unlike the accusative particle, however, the goal/instrumental particle selects *-lo* when the stem ends with a vowel (e.g. *ca-lo*, **ca-ɨlo* 'with a ruler' vs. *ca-lɨl* 'a ruler-ACC') or a liquid (e.g. *khal-lo* 'with a knife' vs. *khal-ɨl*, **khal-lɨl* 'a knife-ACC'). And the allomorph *-ɨlo* is selected when the stem ends with a non-liquid consonant (e.g. *don-ɨlo* 'with money').

As schematically presented in Figure 3, Korean nominative particle system can be viewed from two perspectives: one syntagmatic and the other paradigmatic. It is necessary for a proper allomorph selection mechanism to capture not only the seemingly uninteresting syntagmatic allomorph selection, but also the somewhat inconsistent paradigmatic allomorph selection. As mentioned in Section 2, accounts had to be made under the rule- or constraint-based approaches by resorting to the specification of morphological information, refining the notion of distinctive features, or introducing various complicating constraints. What I propose in this paper is that the noun particle system as a whole can be accounted for by using conditional entropy. Two conditions for the application of the conditional entropy are that the allomorph selection mechanism needs to know what kind of noun particle system is under consideration, and what is the segmental status of the stem final segments. Given these two conditions are given, no other information is needed for selecting correct allomorphs in the Korean noun particle system[9]. If we calculate the total

9 There are cases where the syntagmatic allomorphs can be more complicated than the noun particle systems as shown in the paper, as pointed out by a reviewer.

entropy of the noun particle system from the paradigmatic point of view, we will get 0 as we did for each of the syntagmatic case.

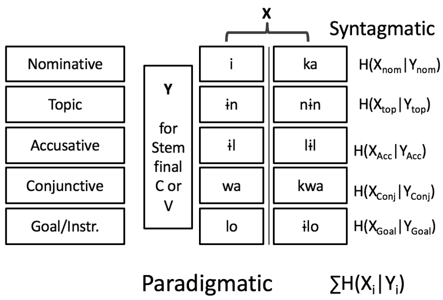

Figure 3. Schematic overview of conditional entropy of each nominative particle allomorphs (referred to as Syntagmatic conditional entropy) and that of nominative particle allomorphs as a whole (referred to as Paradigmatic conditional entropy)

This approach does not require other than the status of the final segment, unlike other rule- or constraint-based approaches, in which feature was specified for a particular segment or constraints were multiplied or special constraint with implicit adjustment of syllables had to be introduced. Furthermore, in this view, it can be entertained that phonological optimization fails to account for the allomorph selection in the Korean noun particles in a consistent way, whereas the conditional entropy can handle those cases in a uniform way. Thus, it is likely that the phonological optimization seen in a subset of the noun particle system (i.e. Nom., Topic, and Acc.) may be a just by-product that has arisen in the minimization process of entropy or uncertainty in the Korean nominative particle system.

4. Discussions and Conclusion

In this paper, I tried to demonstrate the use of quantitative method in suppletive allomorph selection, which has been approached from the perspective of phonology proper, morphology proper, or the interplay of phonology and morphology.

In the tradition of generative phonology, the role of quantitative methods has been mostly neglected. It is evidence by Chomsky's position, which states that "no theory of linguistic structure based exclusively on Markov process models and the like, will be able to explain or account for the ability of a speaker of English to produce and understand new sentences, while he rejects other new sequences as not belonging to the language." (Chomsky 1957). An argument against the statistical methods is partly linked to the argument for the poverty of the stimulus. In this regard, Gold (1967) demonstrated that formal languages cannot be learned without negative evidence. Moreover, negative evidence is not readily available to children. Together, these two facts are widely used as evidence that language is special and largely innate, a line of reasoning known as the "argument from the poverty of the stimulus." However, more recently Manning (2003) outlines evidence that challenges this argument – most importantly, evidence that, unlike categorical grammars, probabilistic grammars are learnable from positive evidence alone (Bod et al. 2003: 6).

Since positive evidence alone can be used as an input to probability grammars, there is a need to integrate probabilities into linguistics. The domain in linguistics which probabilities can be integrated into is "everywhere," according to Bod et al. (2003: 7). Bod et al. (2003) contains papers arguing for the need of probabilities in acquisition, perception and production. It is not merely a tool for processing, linguistic representations, linguistic constraints and well-formedness are also probabilistic. Moreover, probabilistic linguistics does not abandon all the progress made by linguistics thus far; on the contrary, it integrates this knowledge with a probabilistic perspective. The contribution of the current paper is that the quantitative method using conditional entropy can be used to account for the selection of suppletive allomorph.

Entropy seems to have some bearing in our cognition, too. Psycholinguistic modeling of word segmentation also seems to be in support of the use of entropy. Brent and Cartwright (1996) investigated the effectiveness of what they call DR (Distributional Regularity) functions strategies for segmenting broad phonetic transcripts of child-directed English. The DR function in Brent and Cartwright (1996: 107) contains entropy of the relative frequency of the words in the segmentation. Brent and Cartwright (1996) demonstrated that without the component of entropy the DR function did not function well enough for the word segmentation in their experiments. The DR function is defined as in (18):

(18) The DR function defined in Brent and Cartwright (2006)

$$f(S) = 3|\mathbf{TYPES}(S)| + (\log_2 P)\left(\sum_{w \in \mathbf{TYPES}(S)} \ell(w)\right) + |\mathbf{TOKENS}(S)| \times H(S)$$

In (18), S denotes a segmentation, TOKEN(S) is the set of word tokens in segmentation S, vertical bar around a set (e.g., |TOKEN(S)|) indicate the number of items in the set, TYPE(S) is the set of word types in segmentation S, and $l(w)$ is the length of word w, measured in phonemes, and P is the number of phonemes in the input alphabet. And crucially, H(S) is the entropy of the relative frequency of the words in the segmentation.

Now the question is what grammatical component is the suppletive allomorphy computed. The computation may be part of the linguistic component that deals with communicative system, and the resulting partial optimizing effect of phonology might be a by-product of entropy minimization. In this regards, Lee (2015), citing Hume et al. (2011), asserts that "[h]ence we support, without further ado, the fundamentals of the Information Theory in the sense that phonology does not exist for its own sake, and instead, it is the result of optimizing communication." The same line of reasoning is reinforced in Hume et al. (2016), in which the role of phonology is regarded as being "in the service of transferring information." That is, the approach taken in Hume et al. (2016), as in the current paper, is not sound-based but message-based. In the message-based approach, the message typically takes "the form of a morpheme, word, or higher unit of meaning." And the message-based approach predicts that "the shape of phonological patterns occurs as a consequence of the trade-off between resource cost and accurate transmission of these messages."

In conclusion, the approach taken in this paper utilizes the quantifiable measures such as entropy, in particular, conditional entropy. These quantifiable measures successfully induce general patterns that rule or constraint-based approaches fail to capture. Therefore, rule induction through quantifiable measures is effectively used in understanding the full scope of the allomorphic alternation in Korean.

REFERENCES

AHN, SANG-CHEOL. 1996. *An Introduction to Korean Phonology*. Seoul, Korea: Hanshin Publishers.

BOD, RENS, JENNIER HAY AND STEFANIE JANNEDY (eds). 2003. *Probabilistic Linguistics*. Cambridge, M.A.: The MIT Press.

BRENT, MICHAEL AND TIMOTHY CARTWRIGHT. 1996. Distributional regularity and phonotactic constraints are useful for segmentation. *Cognition* 61, 93-125.

BYE, PATRIK 2007. Allomorphy – selection, not optimization. In Sylvia Blaho, Patrik Bye, and Martin Krämer (eds.)., *Freedom of Analysis?,* 63-91. Berlin: Mouton de Gruyter.

CHOMSKY, NOAM. 1957. *Syntactic Structure*. The Hague: Mouton.

EMBICK, DAVID. 2009. *Localism versus Globalism in Morphology and Phonology.* PhD Dissertation. University of Pennsylvania.

GOLD, E. M. 1967. Language identification in the limit. *Information and Control* 10, 447–474.

HALL, KATHLEEN C. 2009. *A probabilistic Model of Phonological Relationships from Contrast to Allophony*. PhD Dissertation. Ohio State University.

HONG, SUNG-HOON. 2014. *Quantitative Methods in P*honology [eum-unlon-ui gyelyangjeog bangbeoblon]. Seoul: Hankwuk Munhwasa.

HUME, ELIZABETH, KATHLEEN C. HALL, ANDREW WEDEL, ADAM USSISHKIN, MARTINE ADDA-DECKER and CEDRIC GENDROT. 2011. Anti-markedness Patterns in French Epenthesis: an Information Theoretic Approach. *Proceedings of the Berkeley Linguistic Society* 37, 1-19.

HUME, ELIZABETH, KATHLEEN C. HALL and ANDREW WEDEL. 2016 Predicting Perceptually Weak and Strong Unmarked Patterns: A Message-based Approach. Vancouver, Canada: 2015 Annual Meeting of Phonology, 9-11 Oct 2015. *Proceedings of the 2015 Annual Meeting of Phonology*.

KAGER, RENÉ. 1996. On affix allomorphy and syllable counting. In URSULA KLEINHENZ (ed.). *Interfaces in Phonology*, 155-171. Berlin: Akademie Verlag.

KIM, GYUN-RRAN. 2006. Suffix-centered allomorphy in Korean. *Studies in Phonetics, Phonology and Morphology* 12.2, 265-281. The Phonology-Morphology Circle of Korea.

KIM, HANSAEM. 2006. Korean national corpus in the 21 century Sejong project. *Proceedings of the 13th National Institute of Japanese Literature (NIJL) International Symposium*, 49-54.

LAPOINTE, STEVENS G. 1999. Stem selection in OT. In Geert Booij and Jaap van Marle (eds.). *Yearbook of Morphology.* 263-297. Dordrecht: Kluwer Academic Publishers.

LEE, PONGHYUNG. 2015. Conditional entropy and deflected contrast for epenthetic vowels in Korean loanwords. *Journal of Studies in Language* 31.2, 377-397.

LEE, YONGSUNG. 2008. Allomorphy in Korean Noun Particles: Multiple Inputs and Default Allomorph. *Korean Linguistics* 14, 65-89.

_____________. 2009. Allomorphy selection: universal and morpheme-specific constraints. *Current Issues in Linguistic Interfaces: Proceedings of The 2009 Seoul International Conference on Linguistic Interfaces*. *Vol. 2*, 417-434.

_____________. 2010. Universal and morpheme-specific constraints for allomorphy selection. *Studies in Phonetics, Phonology and Morphology* 16, 469-490. The Phonology-Morphology Circle of Korea.

ŁUBOWICZ, ANNA. 2006. Opaque allomorphy in OT. In Donald Baumer, David Montero and Michael Scanlon (eds.). *Proceedings of the 25th West Coast Conference on Formal Linguistics*, 261-269. Somerville, MA: Cascadilla Proceedings Project.

MALOUF, ROB. 2009. *Tutorial on maximum entropy models*. Rijksuniversiteit Groningen. [Downloaded from http://www-rohan.sdsu.edu/~malouf/]

MANNING, CHRISTOPHER D. 2003. Probabilistic syntax. *Probabilistic linguistics*, 289-341.

MASCARÓ, JOAN. 2007. External allomorphy and lexical representation. *Linguistic Inquiry* 38.4, 715-735.

MCCARTHY, JOHN AND ALAN PRINCE. 1993. *Prosodic morphology: Constraint interaction and satisfaction*. Ms., University of Massachusetts, Amherst and Rutgers University.

MCCARTHY, JOHN AND MATTHEW WOLF. 2005. Less than zero: Correspondence and the null output. Ms. University of Massachusetts Amherst. [ROA-722]

NEVINS, ANDREW. 2010. Phonologically-conditioned allomorph selection. In Colin Ewen, Elizabeth Hume, Marc van Oostendorp and Keren Rice (eds.). *The Companion to Phonology*. New Jersey: Wiley-Blackwell.

ODDEN, DAVID. 1993. Interaction between modules in lexical phonology. In S. Hargus and E. Kaisse (ed.). *Phonetics and Phonology 4: Studies in Lexical Phonology,* 111–144. San Diego: Academic Press.

OH, MIRA. 1992. Palatal consonants, labial consonants, and vowel epenthesis in Korean loanword phonology. *Linguistic Journal of Korean* 17, 141-161.

PASTER, MARY ELIZABETH. 2006. *Phonological Conditions on Affixation*. PhD Dissertation. The Ohio State University.

________________________. 2008. Phonologically Conditioned Suppletive Allomorphy: Cross-linguistic Results and Theoretical Consequences. In Bernard Tranel (ed.). *Understanding Allomorphy: Perspectives from OT*. London: Equinox.

PATZ, ELIZABTH. 1991. Djabugay. In Robert M.W. Dixon and Barry J. Blake (eds.).

The Handbook of Australian Languages 4, 245-347. Melbourne: Oxford University Press.

PIERCE, JOHN. 1980. *An Introduction to Information Theory: Symbols, Signals and Noise*. New York: Dover Publishers.

SAFFRAN, JENNY, RICHARD ASLIN and ELISSA NEWPORT. 1996. Statistical learning by 8-month-old infants. *Science* 274, 1926-1928.

SHANNON, CLAUDE. E. 1948. *The Mathematical Theory of Communication*. Urbana, IL: University of Illinois Press.

SUH, YUNJU. 2006. Korean Suffix Allomorphy in OT. In *The Proceedings of SUNY/CUNY/NYU Mini-Conference 8.* Linguistics in the Big Apple. [http://web.gc.cuny.edu/dept/lingu/liba/CSN2006.html]

SUNG, EUN-KYUNG. 2005. Allomorph selection in Korean. *Studies in Phonetics, Phonology, and Morphology* 11.1, 45-60. The Phonology-Morphology Circle of Korea.

VAUX, BERT. 1998. *The Phonology of Armenian*. Oxford: Oxford University Press.

YOON, TAE-JIN. 2012. Entropy as a Criterion of Selecting Allomorphs. *Harvard Studies in Korean Linguistics, XIV*. 241-254. Seoul: Hankwuk Munhwasa.

Tae-Jin Yoon
Sungshin Women's University
e-mail: tyoon@sungshin.ac.kr

received: July 12, 2016
revised: November 15, 2016
accepted: December 6, 2016