

Interface between Morphology and Syntax: A Constraint-Based and Lexicalist Approach

Jong-Bok Kim
Kyung Hee University

Jong-Bok Kim. 1998. Interface between Morphology and Syntax: A Constraint-Based and Lexicalist Approach. *Language and Information* 2.1, 177-213. Conflicting criteria used in identifying words have called the lexical integrity principle into question. That is, cases where the morphological word does not coincide with the syntactic word have motivated the syntactic view of word derivation, as pointed out by Bresnan and Mchombo (1995). Further, the implicit desire to make the clausal structure of Korean parallel to those posited for English (Chomsky 1991) and French (Pollock 1989) has also led most of the current literature on Korean morphology to claim that Korean verbal inflections head

There have been several manuscript versions leading to this paper: The first work was presented at the 1st International HPSG Workshop held in 1993 at the Ohio State University. The details of the analysis were greatly changed and a new version was presented at the 2nd International HPSG Conference, University of Copenhagen, Denmark, 1994. Another revised and extensive version has been circulated with the title of 'A Constraint-Based Lexical Approach to Korean verb Inflections' from 1995. This paper, a reduced and revised version of the former, supercedes all these previous ones. Many thanks go to Elizabeth Bratt, Sae-Youn Cho, Chan Chung, Chris Golston, Takao Gunji, Andreas Kathol, Byung-Soo Park, Ivan Sag, Peter Sells, among others for their comments and discussion on the various versions of this paper. I also thank two anonymous reviewers for their suggestions and comments.

their own functional projections such as AgrP, TP, and MP in syntax. In this paper, I will first argue against such a syntactic view. After reviewing some basic properties of Korean verbal inflections, I will show that the evidence from mismatch phenomena supports the lexical integrity principle over the head-movement theories of word derivation. Then, I will propose a theory of lexical grammar which maintains the lexical integrity principle while retaining the effects of functional projections and syntactic movement. (Kyung Hee University)

1. Basic Properties of the Korean Inflection System

1.1. Basic Verbal Suffixes

Korean uses affixation as the major mechanism for word formation. A fully inflected Korean verb consists of a stem, an obligatory suffix and a variety of intervening optional ones. However, the suffixes cannot be freely attached to a stem or word, but have a regular fixed order, as reflected in the traditional template given in (1).

(1) V-root + (Pass/Caus) + (Hon) + (Tense) + Mood

As can be seen in (1), the suffixes marking passive/causative, honorific, tense, and mood information are strictly ordered. A short inventory of these suffixes is given in (2).

- (2) a. Honorific: -si/-usi
 b. Tense: -ess/-ass (Past), -kess (Future)
 c. Mood: -ta (Declarative), -kka (Question), -(e)la (Imperative),
 -(u)psita (Formal Propositive), -ca (Informal Propositive),
 -supnita (Formal Declarative), -supnikka (Formal Question), etc.

Morphologically, the inflectional suffixes preceding Mood are optional, but a Mood suffix needs to be attached to a verb stem occurring in simple sentences. Thus the verbal stem and the mood suffix are mutually bound in the sense that the bare verb stem cannot be used uninflected in any syntactic context, and in that it should be inflected at least mood suffix, as seen in (3).

- (3) a. ilk-(ess)-ta 'read-(Past)-Decl'
 b. *ilk-ess 'read-Past'

1.2. Complementizers

There is another group of verbal suffixes, so-called complementizers, (henceforth COMP suffixes, following Cho and Sells (1995)). They attach not to the final verb in matrix clauses but to the governed main verb in so-called verb-complex constructions. These COMP suffixes can be divided into four groups, according to the kind of stem they can co-occur with, as shown in (4).

- (4) a. COMP1: -a/-e
 b. COMP2: -ci/-key/-ko
 c. COMP3: -eya/-na
 d. COMP4: -ko

These COMP suffixes exclusively occupy the end of a verb form, i.e., no two COMPs can co-occur, and they can be attached only to a specific kind of verb stem:

- (5) a. COMP1: ilk-(*usi)-e 'read-(*Hon)-COMP1'

- b. COMP2: ilk-usi-(**ess*)-ci 'read-Hon-(**Past*)-COMP2'
- c. COMP3: ilk-usi-ess-(**ta*)-eya 'read-Hon-Past-(**Decl*)-COMP3'
- d. COMP4: ilk-usi-ess-(**ta*)-ko 'read-Hon-Past-(**Mood*)-COMP4'

We can observe rigid co-occurrence and combination restrictions here. COMP1 precludes the honorific suffix, COMP2 the tense, and COMP3 the mood suffix. But COMP4 should co-occur with the mood suffixed verb stem.

1.3. Category-Changing Suffixes: Sentential Nominalizers

Unlike the verbal suffixes we have seen so far, there are sentential or inflectional nominalizers such as *-um* and *-ki* that do change the category information of the stem they attach to.

These sentential nominalizers are quite productive, in that there is virtually no restriction to the morphological host, and in that the internal structure of the nominalized clause is fully transparent to other syntactic processes such as scrambling and anaphoric relation, as noted in Yoon (1991).

- (6) a. wuli-nun [John-i hakkyo-ey ka-ss]-UM-ul
 we-TOP John-NOM school-LOC go-Past-Nmlz-ACC
 al-ass-ta.
 know-Past-Decl
 'We knew that John went to school.'
- b. [namca-ka yeca-lul salangha]-KI-ka elyep-ta.
 man-NOM women-acc love-Nmlz-NOM difficult-Decl
 'It is difficult for man to love woman.'

The nominalizers *-um* and *-ki* in (6a,b) attach to the tensed

stem and underived verb root respectively, and they nominalize the embedded clause.

The productivity and position of these nominal suffixes have led linguists to regard them as inflectional suffixes, as reflected in the commonly assumed template in (7).

(7) V-root + (Pass/Caus) + (Hon) + (Tense) + Nmlz

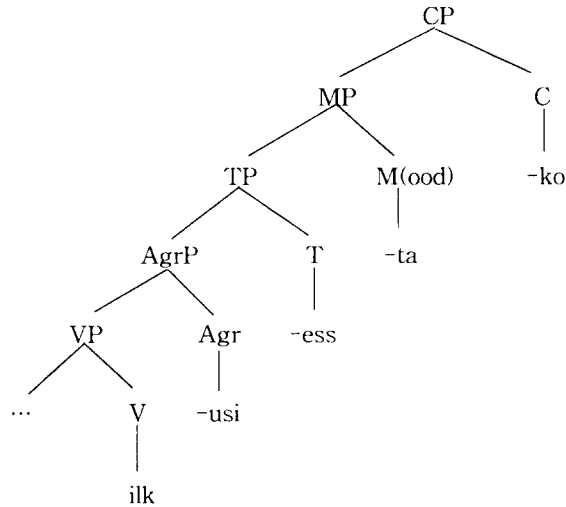
As we have seen so far, each verb suffix in Korean exhibit rigid suffix ordering and combinatory restrictions with respect to other verb suffixes and stems. In the next section, I will examine the syntactic view which attempts to capture part of these properties, such as ordering restrictions among suffixes, by hierarchical functional projections.

2. Review of Syntactic Approaches

2.1. Arguments

Much influenced by Chomsky (1991) and Pollock (1989), most of the current literature on Korean morphology has claimed that Korean verbal inflections are formed in syntax. They commonly break up an inflected verbal word into several morphemes according to their functions and then assign them corresponding syntactic nodes. The representation in (8) is a typical and simplified syntactic tree in deriving the verb *ilk-usi-ess-ta-ko* 'read-Hon-Past-Decl- COMP4'.

(8)



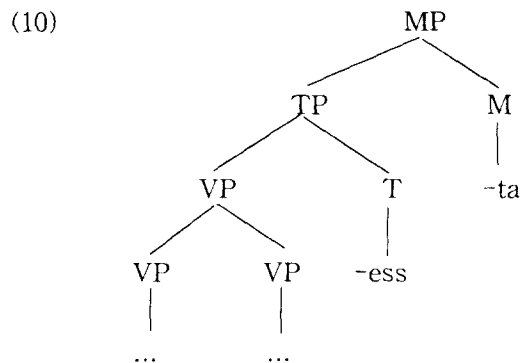
Here, each morpheme occupies a functional category. And, according to head-movement, the verb *ilk* moves to the next higher functional category, Agr, and then the resulting complex morpheme moves up again. By this kind of successive head movement process, the final surface constituency and correct ordering of the morphemes are obtained.

Such a syntactic view which derives the morphemic structure of words by means of syntactic head movement has been basically motivated from mismatch between morphology and syntax (semantics). In Korean, the phrasal distribution of some suffixes and their wide scope have been taken to be strong evidence for the syntactic view.

Let us review some of these basic arguments. First, in coordinate structures, when the tense suffix appears only in the final conjunct, it takes distributive scope over the non-final conjunct which is unmarked for tense.

- (9) [[[[Kim-un pap-ul mek-ko], [Lee-nun ppang-ul mek]]-ess] -ta]
 meal-ACC eat-CONJ bread-ACC eat-Past-Decl
 'Kim ate (*eats) a meal, and Lee ate bread.'

Though the past tense suffix *-ess* in (9) is part of the main verb in the second conjunct, it semantically scopes over the whole sentence. This can be seen by the fact that the tenseless verb in the first conjunct should be interpreted as past tense verb, not as present tense. Under the assumption that only constituents (e.g., VP) can be coordinated, and that the subject is generated VP-internally, the syntactic view claims that sentences like (9) involves a VP coordination, as represented in (10).



The structure in (10) appears to naturally account for the suffix order and the obligatory distributive scope of the tense information.

There are several more cases where we can find a mismatch between morphology and syntax. For example, consider the sentential nominalization case again in (11).

- (11) [[ai-tul-i cip-eyse kongpwuha]-ki]-ka swipci anhta.
 child-PL-NOM house-LOC study-Nmlz-NOM easy not
 'It is not easy for children to study at home.'

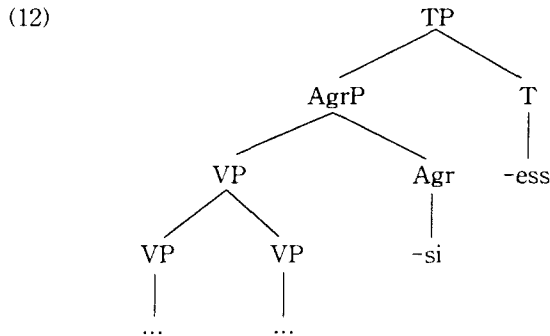
As we have seen before, the sentential nominalizer morphologically forms a word with the predicate *kongpwuha-* in (11). However, the suffix *-ki* syntactically nominalizes the sentential complement and makes it the sentential matrix subject.

2.2. Some Problems

By introducing inflections in syntax, the syntactic view appears to deal with their scope aspects and phrasal distribution naturally. However, when we look into the facts in more detail, it turns out that the syntactic view requires redundant or supplementary ways of capturing the complex behavior of the Korean verbal inflection system.

2.2.1. The Status of AgrP

We have seen that the distributive property of the tense information is accounted for by assuming VP conjunction with the phrasal suffixation of the tense suffix. If we follow the same line of inference with the hierarchical structure in (8) where the honorific suffix has its own projection, Agr, then we predict that the honorific information coming from the Agr should also be able to distribute over non-final conjuncts, as can be seen from the structure (12).



There is no reason to block this structure in the syntactic view where each suffix has its own functional projection.

However, when the final conjunct is marked with the honorific information but non-final conjuncts are not, the honorific information should not distribute over non-final conjuncts. This can be seen from the sentence in (13).

- (13) [haksayng-tul-un paywu-(*si)-ko], [sensayng-nim-un
 student-PL-TOP learn-(*Hon)-CONJ teacher-Hon-TOP
 kaluchi-si-ess-ta]
 teach-Hon-Past-Decl
 'The students were learning, and teachers were teaching.'

The honorific verb suffix *-si* agrees with the subject that carries the honorific nominal suffix, *-nim*. This is the case in the final conjunct in (13). However, in the first conjunct, the presence of the honorific suffix in the verb *paywu-* 'learn' will violate the agreement condition since the subject is not honorific-marked. If we assume the structure (10) and follow the same explanation for the tense distribution in the syntactic view, then we predict that the agreement information also should

distribute over the first conjunct. To prevent the honorific information of the final conjunct from overriding that of the non-final conjunct, the syntactic view thus needs to say that unlike tense and mood suffixes, the honorific suffix does not form an independent functional projection.

There is another argument showing that there should be no Agr functional head in the syntax. Though one of the main arguments for functional projections is based on the productivity of all the verb inflectional suffixes, the honorific suffix *-si* is not totally productive. There are cases that show lexical idiosyncrasies with respect to the honorific suffix. The roots such as *mosi-* 'serve', *cwumusi-* 'sleep' in (14), inherently have the honorific suffix *-si*, while the verb root like *mek-* 'eat' can never occur with it. This means that these verb roots are lexically specified with the positive (*-si*) or negative (\emptyset) honorific suffix, as also noted in Han (1991). However, the syntactic view, where the functional head, Agr, selects for a VP headed by the verb root, may allow us to generate ill-formed combinations such as given in (14).

- (14) a. *cwumusi-(*si)-ta* 'sleep(Hon)-*Hon-Decl'
 b. *mosi-(*si)-ta* 'serve(Hon)-Hon-Decl'
 c. *mek-(*usi)-ta* 'eat-(*Hon)-Decl'

The syntactic view predicts that these are well-formed combinations since the honorific suffix occupies the Agr and it is a head in syntax. This further supports the claim that *-si* does not form an independent projection.

One possible way of explaining the impossibility of the double marking here is to stipulate that the Agr head *-si* cannot select

for a VP whose head is marked with the feature [+Hon], as also noted by Yoon (1993). Even though this solution somehow works out, there exists another argument showing that *capswu-si* is lexically one inseparable unit, as noted in Han (1991). We have seen that COMP1 *-e* can combine only with a non-honorific suffixed verb stem. But the irregular honorific verb *capswu-si* can occur with the COMP1 suffix whereas the regular honorific verb *ilk-usi* cannot, as in (15).

- (15) a. *capswusi-e* (po-si-ess-ta)
 eat(Hon)-COMP1 try-Hon-Past-Decl
 'tried eating'
- b. *ilk-(*usi)-e* (po-si-ess-ta)
 read-Hon-COMP1 try-Hon-Past-Decl
 'tried to read'

What the contrast tells is that *capswusi* and *ilk-* are both the same kind of morphological unit, verb root. If we assume the honorific suffix as an independent projection, it is hard to account for this contrast. But the lexical view, where all verbs are lexically specified for the honorific information either suppletively or by productive morphemes, can block the double marking simply by a general morphological blocking principle such as No Vacuous Suffixation Principle.

2.2.2. Tense Distribution in Coordination

Now let us consider tense distribution. The first question that arises is whether or not examples like (10) where the non-final conjunct is untensed are really sentential coordinations. There is

some evidence showing that the alleged non-tensed conjunct is a subordinate clause, and thus the tenseless conjunction is not a real coordination, but an adjunction.

First, the occurrence of the events described by the untensed clause and the final clause must be sequential or at least parallel. Consider the sentences given in (16).

- (16) a. Kim-i mence sangca-ey chayk-ul neh-ko Lee-ka
 Kim-Nom first box-Loc book-ACC put-CONJ Lee-NOM
 kukes-ul nacwungey phocangha-yess-ta.
 that. thing-ACC later wrap-Past-Decl
- b. *??Kim-i nacwung-ey sangca-ey chayk-ul neh-ko Lee-ka
 kukes-ul mence phocangha-yess-ta.

The sentence (16a) where the first clause is untensed is taken to be a VP coordination (Yoon 1991) having a structure similar to the one given in (12). But notice here that the contrast between (16a) and (16b) shows that the ordering of the temporal adverbs *mence* 'first' and *nacwungey* 'later' in each assumed conjunct cannot be violated.

An interesting point is that in IP coordination (in Yoon's (1993) term) where all conjuncts are tensed the reverse order of the two temporal adverbs is possible, as can be seen in (17).

- (17) Kim-i nacwungey sangca-ey chayk-ul neh-ess-ko Lee-ka
 Kim-Nom later box-Loc book-ACC put-Past-CONJ Lee-NOM
 ku kes-ul mence phocangha-yess-ta.
 that thing-ACC first wrap-Past-Decl

I will argue that this contrast comes from the fact that the

-ko conjunction marking in (16b) and (17) are not identical: there are at least two different usages of *-ko*. One is the real conjunction marking corresponding to English ‘and’, and the other is a temporal-sequential marking similar to English ‘and then’ or ‘after’. And further, I will show that the untensed clause behaves like a modifying subordinate clause.

One strong piece of evidence supporting the assumption that the non-tensed clause is a modifier clause comes from an asymmetry in anaphor binding. In Korean, it has been observed that the anaphor *caki* in a subordinate clause can take the subject of the main clause as its antecedent. Then consider the pair in (18).

- (18) a. *caki_i-uy atul-i colupha-ko Kim_i-i kapcaki*
 self-GEN son-NOM graduate-CONJ Kim-NOM suddenly
cwuk-ess-ta.
 die-Past-Decl.
 ‘After self_i’s son graduated, Kim_i died suddenly.’
- b. **caki_i-uy atul-i colupha-yess-ko Kim_i-i kapcaki*
cwuk-ess-ta.

In the tenseless case (18a), the anaphor can be bound by the subject of the main clause, whereas in the tensed case (18b), it cannot. If it is true that (18a,b) are both VP/IP coordination, the difference in the anaphor binding cannot be captured easily.

Case alternation facts also support the assumption that non-tensed clause is an adjunct. It is a well-known fact that desiderative predicates like *siph-* ‘want’ allow Nom/Acc alternations as shown in (19).

- (19) Kim-un maykcwu-ka/lul masi-ko siph-ess-ta.
 Kim-TOP beer-Nom/Acc drink-COMP want-Past-Decl
 'Kim wanted to drink beer.'

But, consider the examples in (20).

- (20) a. ??/*Kim-un [maykcwu-ka masi-ko] [pap-ul mek-ko]
 Kim-TOP beer-Nom drink-CONJ meal-ACC eat
 siph-ess-ta.
 want-Past-Decl
 'Kim wanted to drink beer and eat the meal.'
- b. ?Kim-un maykcwu-lul masi-ko pap-i mek-ko siph-ess-ta.
 c. Kim-un maykcwu-lul masi-ko pap-ul mek-ko siph-ess-ta.

Of these sentences, (20a) especially shows that in contrast to the final clause, no case alternation is allowed in the first non-tensed clause. If we accept the view that sentences in (20) are VP coordinations, then we may need an additional mechanism or stipulation to account for why the object of the first conjunct does not undergo case alternations, whereas that of the final does. However, provided that the first non-tensed clause is not a conjunct but an adverb phrase which modifies the second clause, this fact can easily be accounted for: the predicate *sip-ess-ta* 'wanted' does not affect the case marking of the complements in the modifier clause.

I have claimed that while the coordination of two tensed clauses is a real coordination, that of untensed non-final conjuncts with the tensed final one is an adjunction. Then it is expected that only the former cases may show the Coordination Structure Constraint. This prediction is borne out from the difference in the extractability of a complement between the untensed conjunct and tensed one.

- (21) a. [Tom-i pap-ul mek-ko] [Mary-ka selkeci-lul
Kim-Nom meal-Acc eat-CONJ Mary-Nom clean-Acc
ha-yess-ta]
do-Past-Decl
'Tom ate the meals and Mary washed the dishes.'
- b. Mary-ka [Tom-i pap-ul mek-ko] [t selkeci-lul ha-yess-ta]
c. selkeci-lul [Tom-i pap-ul mek-ko] [Mary-ka t ha-yess-ta]
d. Mary-ka selkeci-lul [Tom-i pap-ul mek-ko] [t t ha-yess-ta]

As can be observed in (21b,c,d), the element(s) in the second tensed clause can freely be extracted. However, a similar test in (22) shows that the real coordination where the verb in each conjunct is tensed does not allow any element to be extracted, observing the Coordination Structure Constraint.

- (22) a. [Tom-i pap-ul mek-ess-ko] [Mary-ka selkeci-lul
Tom-Nom meal-Acc eat-ess-CONJ Mary-Nom clean-Acc
ha-yess-ta]
do-Past-Decl
'Tom ate the meals and Mary washed the dishes.'
- b. *Mary-ka [Tom-i pap-ul mek-ess-ko] [t selkeci-lul
ha-yess-ta]
c. *selkeci-lul [Tom-i pap-ul mek-ess-ko] [Mary-ka t
ha-yess-ta]
d. *Mary-ka selkeci-lul [Tom-i pap-ul mek-ess-ko] [t t
ha-yess-ta]

The extraction of the element(s) in the final clause is possible when the main verb of the first clause is untensed. Such an extraction is, however, prohibited when it is tensed. Considering that Korean generally observes the Coordinate Structure Constraint the possibility of extracting elements in the non-tensed coordination further supports that it is not a real

coordination, but a modifying adverb clause.

Given these arguments for assuming the untensed clause as an adjunct, we can easily account for the tense dependency in examples like (9). I will briefly explain how we can do this. At first, adopting Pollard and Sag's (1994) approach wherein adjuncts are functions that take heads as arguments, we can represent the lexical entry for the untensed *-ko* suffixed verb as in (23).

- (23)
$$\left[\begin{array}{l} \text{CAT|HEAD } \textit{verb} \text{ [MOD VP[LOC } \boxed{3}:\boxed{2}] \\ \text{CONT } \left[\begin{array}{l} \text{RELATION } \textit{and-then} \\ \text{ARG1 } \boxed{1} \\ \text{ARG2 } \boxed{2} \\ \text{LOC(ATION) } \boxed{3} \end{array} \right] \end{array} \right]$$

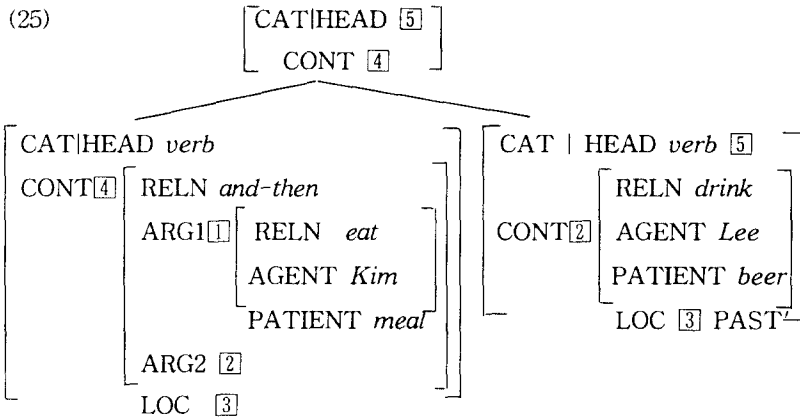
This lexical entry says that the *-ko* attached verb functions as a modifier to the V-projection (indicated as MOD feature). And it further says that the suffix *-ko* semantically places the semantic relation *and-then* between the untensed and tense clause.

Given this lexical entry, we then are able to obtain the correct tense distribution for the sentence (9), repeated here as in (24).

- (24) [[[[Kim-un pap-ul mek-ko], [Lee-nun maykcwu-ul
Kim-TOP meal-ACC eat-CONJ Lee-TOP beer-ACC
masi]]-ess]-ta]
drink-Past-Decl
'Kim ate (*eats) a meal, and Lee drank beer.'

According to our system, this sentence will have the simplified structure (25), represented within the grammar of Head-driven Phrase Structure Grammar.'

⁴For details of the grammar, readers are referred to Pollard and Sag (1994).



As specified in the lexical entry of the *-ko* suffixed verb in (23), the semantic relation *and-then* holds between the semantic content (ARG1 [1]) of the first untensed clause and that of the head (ARG2 [2]) it modifies. Thus, when the untensed and final clauses are combined, we obtain the reading of the sentence (24), informally represented as in (26).

- (26) PAST' *and-then* $\langle eat'(k,m), drink'(l,b) \rangle$ where PAST' denotes a time, t' , that precedes the utterance time, t .

2.2.3. Summary of the Issues in the Syntactic View

So far, we have seen that the behavior of the honorific suffix and the adjunct property of the untensed clause crucially undermine the validity of the syntactic view. In what follows, I will propose a lexical approach, built upon the constraint-based framework of HPSG.

3. A Constraint-Based and Lexicalist Approach

Contrary to the syntactic view which assumes that word formation occurs both in the syntax (mainly inflections) and in the lexicon, the lexical view claims that word formation occurs only in the lexicon. In the lexical view, suffixes (both inflectional and derivational) exist in the domain of the lexicon, an independent component of the grammar. Such a lexical view has been developed on the principle of lexical integrity stating that the syntax neither manipulates nor has access to the internal form of words. Then, the issue in the lexical view is how to formalize a more efficient lexical structure which can enrich the lexical information and thus account for morphological phenomena to which syntax seems to make direct or indirect reference.

In what follows, I will lay out a lexical view and show how this view can deal with such issues, under the framework of HPSG which is developed by Pollard and Sag (1987, 1994), and many others. In HPSG, where feature structures serve as models of linguistic objects, Pollard and Sag (1987) have successfully developed two mechanisms: sort hierarchies and lexical rules to eliminate massive redundancies ('vertical and horizontal')³ caused in word formation. The fundamental reason for adopting the mechanism of 'sort hierarchy' in the lexical structure is to provide a way to minimize the information for each lexical entry.

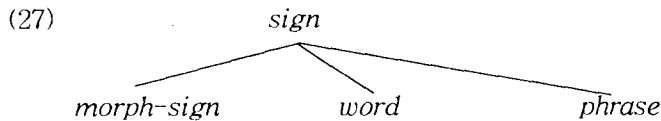
³Pollard and Sag (1987) classify two kinds of redundancy: 'vertical' redundancy is the one caused by encoding in each lexical entry all its linguistic information shared with whole word classes, such as parts of speech, valence, classes, and so forth. And 'horizontal' redundancy is the one that comes from groups of words whose specific information contents are correlated according to recurrent patterns, such as in inflectional and derivational paradigms.

Though HPSG analyses have successfully developed hierarchical lexical structures and lexical rules to eliminate all redundancy from lexical descriptions, my analysis of Korean morphology makes use of only ‘sort hierarchy’ and ‘definitional constraints’. I will show that the system can capture the combinatory and ordering generalizations of verbal suffixes with no additional machinery. This analysis will also give us ways to account for interactive phenomena between morphology and syntax, without violating the principle of lexical integrity.

3.1. The Notion of *Sort*

Let us first consider in more detail some key notions on sort hierarchy and feature structures relevant for morphology.

In HPSG which adopts formalized feature logic, feature structures are all *sorted*. The set of all sort symbols is defined to be partially ordered, and they construct a sort hierarchy. For example, the basic linguistic object in HPSG is the sort, *sign*. This basic sort can be declaratively partitioned into three subsorts, *morph-sign*, *word* and *phrase*:



Objects of *morph-sign* are morphological objects that have word internal structures. In addition to such declarations on the sorts, we need another kind of declaration in order to formulate the sort system into the grammar. It is an appropriateness condition stating that all feature structures should be *well-typed*.

This means that a *sort* determines what kind of attribute or feature label the feature structure of the sort can have. The mechanism of feature declarations will ensure that each well-defined sort is associated with a specific feature structure. For instance, the sort, *sign* is declared to have the feature structure given in (27). This says that objects of *sign* must have the attribute PHON(OLOGY) whose value is defined to be of the sort *list* and SYNSEM (syntax and semantics) whose value in turn is of the sort *synsem*.

$$(28) \left[\begin{array}{ll} \textit{sign} & \\ \text{PHON} & \textit{list} \\ \text{SYNSEM} & \textit{synsem} \end{array} \right]$$

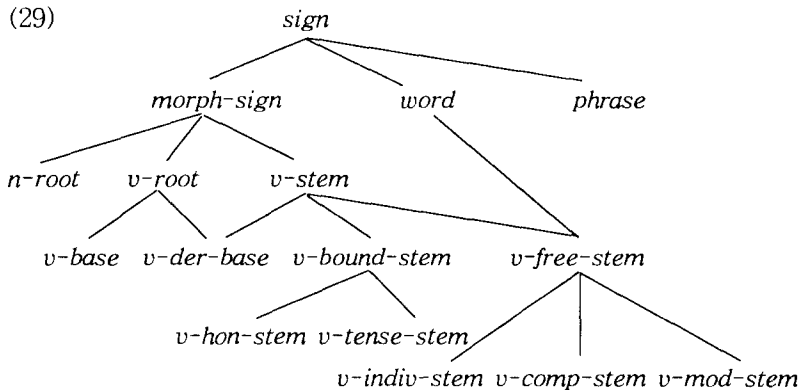
These two mechanisms, sort hierarchy and feature declarations, predict that a sort will inherit the feature declarations of its supersorts. For example, since *morph-sign* is a subtype of *sign*, this sort will inherit all the constraints that its supersort *sign* carries, in addition to its own constraint that it has STEM value. This in turn means that linguistic objects of the sort, *morph-sign* will minimally have PHON and SYNSEM attributes, inherited from its supersort *sign*. Such a mechanism provides ways of eliminating the ('vertical and horizontal') redundancies in word formation: since sorts (or types) in a given hierarchy hold subsumption relations,⁴ all the lexical information we need to encode in the lexicon is the information that is not inherited from the supersorts of that lexical element.

⁴A sort A is subsumed by another sort B if A is at least as informative as B. Thus if a subsumption relation holds between two sorts, then a subsumed sort (subsort) will inherit all the information that its supersorts bear.

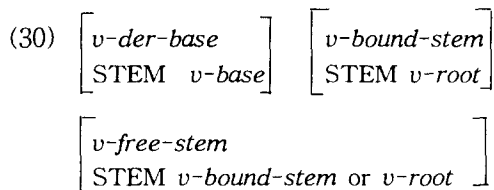
3.2. Sort Hierarchy and Constraints for Korean Morphology

3.2.1. Basic Sort Hierarchy

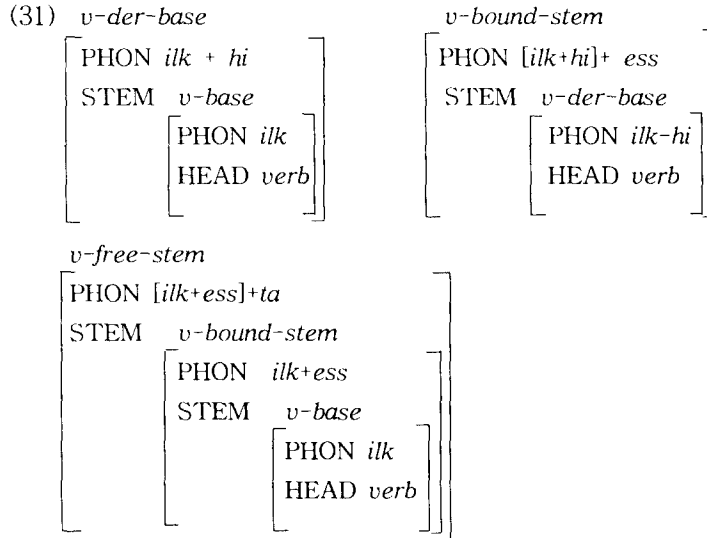
Given this background, I assume the following simplified sort hierarchy for Korean, constructed upon the basis of the morphological behavior of verbal suffixes in Korean:



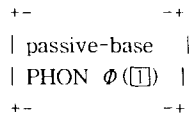
The basic linguistic sort, *sign*, is declared to have three subsorts, *morph-sign*, *word*, and *phrase*, each of which has its own ontological value. Further for a sort to be *well-typed*, each sort has its own feature values either by inheritance or by its own constraint. For example, the sort *morph-sign* is declared to have the feature STEM. However, its subsorts will have their own constraints:



The constraints in (30) restrict the possible word internal structures in Korean word formation. Structures in (31) are examples of the possible morphological structures that such declarative constraints on the STEM value of each subsort allow:



*The suffix value of each stem is realized by a functional operation. For example, the sort passive-base will have the following function for its suffix value.



where $\emptyset(\square)$ is defined as follows:

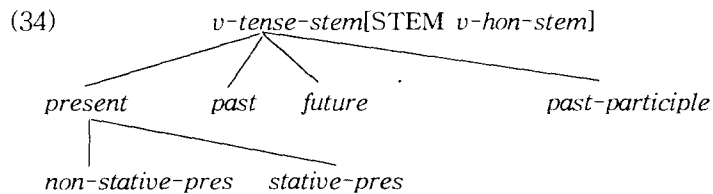
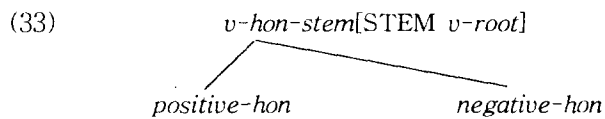
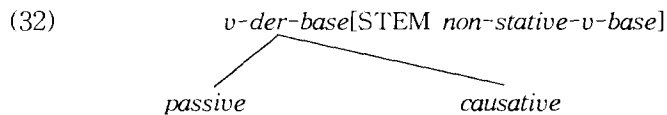
- \square + -i/([-consonant], [-continuant, +tense])__ (e.g., noh-i-ta 'locate-Pass-Decl')
- \square + -hi/([-continuant, -tense])__ (e.g., mek-hi-ta 'eat-Pass-Decl')
- \square + -li/[+lateral]__ (e.g., mol-li-ta 'drive-Pass-Decl')
- \square + -ki/[+consonant, -plosive, -lateral]__ (e.g., an-ki-ta 'hug-Pass-Decl')

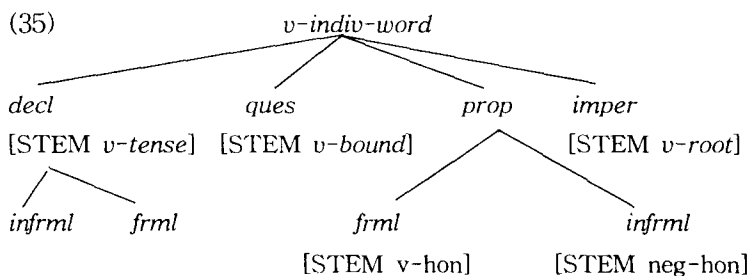
We leave open defining functions for how all the verbal suffixes are realized depending on their phonological environment.

Such a constraint system provides a clean account for the ill-formed combinations of morphological elements without employing an additional mechanism such as morphological ordering templates. For example, the *v-der-base* **[ilk+ess]+hi* is ill-formed simply because it violates the constraint that its STEM value be *v-base*. In the same vein, **[ilk+ess+ta]+ess* is unacceptable because its STEM value is not *v-der-base* but *v-free-stem*.

3.2.2. Further Partitions and Constraints on Verbal Sorts and Their Implications

In order to deal with strict ordering and combinatory possibilities among suffixes, we need further classification for the subsorts in the hierarchy (29):





Let us consider a few relevant points.

First, the detailed specifications on the STEM's value on certain sorts are necessary to deal with ordering and co-occurrence restrictions between suffixes.

- (36) a. *[[cap-ass]-si]-ta 'catch-Past-Hon-Decl'
 b. *[[cap-usi]-ta]-ess 'catch-Hon-Decl-Past'
 c. *[[[cap-usi]-usi]-ess]-ta 'catch-Hon-Hon-Past-Decl'
 d. *[[[cap-usi]-ess]-ta]-ta 'catch-Hon-Past-Decl-Decl'

The verb forms in (36) are unacceptable combinations because of the incorrect ordering of verb suffixes, and unwarranted recursive occurrence of a suffix. All these data can be explained without any ordering rules between suffixes.

For example, (36a) is ruled out because the honorific suffix co-occurs with the *v-tense* stem, *cap-ass*. As can be seen in (33), the only possible stem it can co-occur with is a *v-root*.

In the hierarchy, each sort has its own STEM value, either inherited from its supersorts or declared as its own property. At first glance, the constraints on the STEM value given to specific sorts may appear equivalent to assigning strict subcategorization information to each suffix. However, there is one important difference. We do not specify co-occurrence

constraints on each sort. We factor out regularities and irregularities in co-occurrence restrictions. For example, the sort, *v-hon* has its own STEM value [STEM *v-root*]. This constraint should be observed by its subsorts due to the basic inheritance in the sort hierarchy. Such a system greatly reduces the (vertical and horizontal) redundancy since on my approach, where a group of morphological objects that share information such as ordering and combinatory constraints, is sorted out, we need to state the relevant constraint on this localized group (sort) just once.

One important question arises: why do we need the notions of sorts in the morphological theory? The reason is simply that any morphological theory for Korean needs certain notions similar to sorts. We can find cases where we should have some notions referring to a specific group of morphological objects, so as to predict that a certain morphological phenomenon applies only to this group. As seen in (35), the sort *infrml prop (informal propositive)* requires STEM's value to be *negative-hon*. Consider the data in (37).

- (37) a. *ilk-usi-ca 'read-Hon-Prop'
 b. *cwumusi-ca 'sleep(Hon)-Prop'
 c. ilk-ca 'read-Prop'
 d. ca-ca 'sleep-Prop'

(37a) and (37b) show that the suffix *-ca* cannot occur either with a regular honorific form (*ilk-usi*) or with a suppletive honorific form (*cwumusi*). It occurs only with a verb stem with negative honorific information, as can be seen from (37c) and (37d). To capture such co-occurrence restrictions, any theory

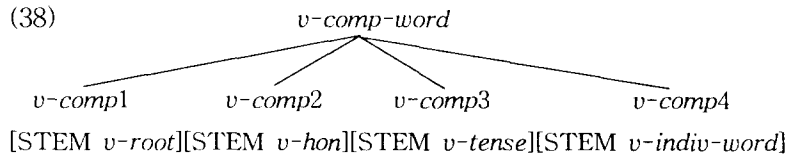
needs to refer to the unit of *negative-hon*. These facts tell us why we need a notion (*sort* in our system) that can refer to a specific group of morphological objects. Each sort defined here has its own reason to exist theory-independently.

3.2.3. Objects of Free-Stem

In the rest of this paper, I will show how the constraint-based approach proposed here can further describe some interactive phenomena between morphology and syntax.

3.2.3.1. Complementizers

We have seen in Section 1 there are four kinds of COMP suffixes, according to the kind of stem they can co-occur with. Given their behavior, the sort *v-comp-word* can be partitioned into four subsorts:



The four subsorts of *v-comp-word* are declared to have their own feature values for STEM. The declarations on the STEM value of each sort given in (38) will correctly generate only well-formed *v-comp-word* elements. Consider the examples in (39).

- (39) a. *[cap-ass]-a (po-ta)
 catch-Past-COMP1 try-Decl

- 'try to catch'
 b. *[cap-ass]-ci (anh-ass-ta)
 catch-Past-COMP2 not-Past-Decl
 'did not catch'

The ill-formedness of (39) is correctly predicted. For example, the generation of (39a) is blocked since the STEM's value for the *v-comp1* can be only a *v-root*, but *cap-ass* is of the sort *v-tense*.

Another related issue in the COMP suffixes is that when they appear on main verbs in verb complex, their form is determined by the following auxiliary verb.

- (40) a. *cek-e anh-ass-ta.
 write-COMP1 not-Past-Decl
 'did not write'
 b. cek-usi-ci anh-assta.
 write-Hon-COMP2 not-Past-Decl
 'did not write honorably'
- (41) a. cek-e po-ass-ta
 write-COMP1 try-Past-Decl
 'tried to write'
 b. *cek-usi-ci po-ass-ta
 write-Hon-COMP2 try-Past-Decl
 'tried to write honorably'

The ungrammaticality of (40a) and (41b) is due to the fact that the second auxiliary verb does not combine with the main verb with the proper COMP form: the negation Aux in (40) requires the main verb with a COMP2 suffix, and the equi auxiliary *try* in (41) takes a main verb with the COMP1 suffix. This constraint applies to all Aux verbs. To account for such a generalization, all we need to do is to define a local

constraint to the sort *v-comp-word*, as in (42).

$$(42) \left[\begin{array}{l} v\text{-comp-word} \\ \text{PHON } \emptyset ([1]) = [1] + [2] \\ \text{SYNSEM|CAT|HEAD } verb \text{ [VFORM } [2]] \\ \text{STEM|PHON } [1] \end{array} \right]$$

The feature description in (42) says that objects of the sort *v-comp-word*, have the constraint that its VFORM value is identical to the affix's PHON value mapped from the function \emptyset .

This general constraint makes it possible to capture the generalization that COMP suffixed verbs exhibit.

For example, consider the morphological structure of *cap-a* 'catch-COMP1' in (43).

$$(43) \left[\begin{array}{l} \text{PHON } [1]cap + [2]a \\ \text{SYNSEM|CAT|HEAD } verb \text{ [VFORM } [2]a] \\ \text{STEM PHON } [1]cap \end{array} \right]$$

Given this, all we need to explain complex predicate data in (40)-(41) is now the subcategorization information for the second auxiliary verb.

- (44) a. *ahn-* 'not': SUBCAT < ... V[VFORM *ci*] ... >
 b. *po-* 'try': SUBCAT < ... V[VFORM *a*] ... >

Together with the local constraint given in (42), the lexical entries in (44) will correctly predict the data given in (40) and (41). The negation verb *ahn-* 'not' selects for a verb with the VFORM value *-ci* to form a complex predicate, while the verb *po-* 'try' requires a verb with the VFORM value *-a*, all of

which become visible at the level of the morphological mother due to the local constraint in (42).

Such an analysis also gives an answer to the issue on the right-headness, which has been problematic for syntactic approaches, as also noted in Sells (1995). Let us consider rather a complicated complex predicate as in (45).

- (45) *cap-a/*ci-man-un* *po-ass-ta*
 *catch-COMP1/*COMP2-Del-TOP try-Past-Decl*
 'tried to catch at least ...'

The Aux *po-* can combine only with a COMP1 suffixed verb, as previously noted. In the syntactic view where a verbal suffix is taken to be a head, such an example violates the locality or adjacency of selection. If we simply assume the delimiters *-man* and *-un* are independent functional projections, then the information about the form value of COMP suffixes would not be passed up to the resulting form *cap-a-man-un*. Thus the syntactic view needs to say either that the Aux *po-* needs to look several layers down into the structure of the governed main verb in order to see its COMP form.

In our approach where suffixes neither are functional heads nor exist as independent syntactic projections, such an issue does not arise. All we need to say is that a delimiter or topic marker after a *v-comp* word does not register VFORM value on its morphological mother. The final morphological structure of *cap-a-man-un* in (46) that our system generates correctly encodes the VFORM value without violating the locality condition.

- (46) $\left[\begin{array}{l} \text{PHON } cap+a+man+un \\ \text{HEAD } verb[VFORM a] \\ \text{STEM } \left[\begin{array}{l} \text{PHON } cap+a+man \\ \text{HEAD } verb[VFORM a] \\ \text{STEM } \left[\begin{array}{l} \text{PHON } cap+a \\ \text{HEAD } verb[VFORM a] \end{array} \right] \end{array} \right] \end{array} \right]$

3.2.3.2. Sentential Nominalizations

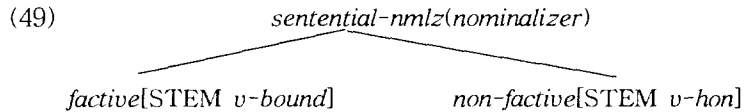
We have seen that there are two main sentential nominalizers, *-um* and *-ki*. There are some semantic distinctions between *-um* and *-ki*. The former has been said to be a factive, and the latter to be a nonfactive nominalizer, as shown in (47).

- (47) a. Tom-un [kunye-ka ttena-ass-*ki/um]-ul hwuhoyha-n-ta
 Tom-TOP she-NOM leave-Past-Nmlz-ACC regret-Pres-Decl
 'Tom regrets that she left.'
- b. Tom-un [kunye-ka ttena-(**ass*)-ki/*um]-lul kitali-ess-ta.
 Tom-TOP she-NOM leave-Past-Nmlz-ACC wait-Past-Decl
 'Tom waited for her leaving.'

Sentences in (47) show another difference between *-um* and *-ki*. Morphologically, *-um* can occur with a tense suffixed stem, but *-ki* cannot. There is also another important difference. Only *-um* can be used as a kind of mood marker, marking the sentence ending.

- (48) a. ku-ka ttena-ess-um
 he-NOM leave-Past-Nmlz
 'His having been left.'
- b. *ku-ka ttena-ki

Given these differences which support the existence of two different morphological objects with respect to sentential nominalizations, we can construct the hierarchy as given in (49).

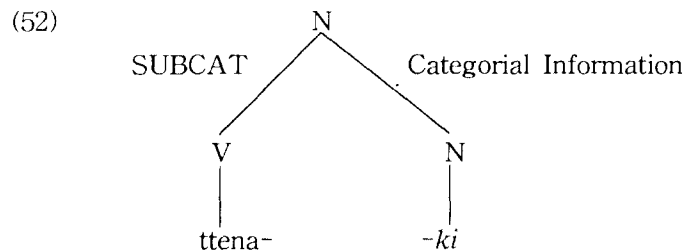


The constraints of the STEM's value on each sort can account for the data given in (50) and (51).

- (50) a. [ilk-(usi)-(ess)]-um 'read-Hon-Past-Nmlz'
 b. *[ilk-usi-ess-ta]-um 'read-Hon-Past-Decl-Nmlz'
- (51) a. *[ilk-ess]-ki-lul (hayessta) 'read--Past-Nmlz-ACC (did)'
 b. [ilk]-ki-lul (hayessta) 'read-Nmlz-ACC (did)'

(50a) is well-formed since *-um* combines with a *tense-stem* which is a subsort of *v-bound*. But (50b) is ill-formed since the nominalizer *-um* occurs with a mood (Decl) suffixed *v-indiv-word*. (51a) is ruled out since *-ki* cannot occur with a *tense-stem* object, according to our constraint.

Another thing to note in sentential nominalizations is that there is an issue in information flow.



As can be seen in (52), the SUBCAT information is coming from the left element since the nominalized N still needs to combine with the complement(s) of the verb, while the categorial information comes from the righthand nominalizer. This kind of information conflict is not problematic in our approach where suffixes are no longer heads. All we need to do is to give a proper constraint to the morphological object, the sort *sentential-nmlz* as given in (53).⁶

(53) Constraint on the sort *sentential-nmlz*:

PHON \emptyset (3) HEAD <i>noun</i> VAL [1] CONT [2] STEM [<table style="border-collapse: collapse; display: inline-table; vertical-align: middle;"> <tr> <td style="border-left: 1px solid black; border-right: 1px solid black; padding: 2px 5px;"> PHON [3] HEAD <i>verb</i> VAL [1] CONT [2] </td> </tr> </table>]	PHON [3] HEAD <i>verb</i> VAL [1] CONT [2]
PHON [3] HEAD <i>verb</i> VAL [1] CONT [2]	

This constraint on the sort *sentential-nmlz* specifies that the categorial information of its STEM should be changed into a noun. This is the main property that both *-ki* and *-um* nominalizers contribute to the resulting form. This lexical specification avoids the issue of the information conflict, and hence allows us to obtain the proper information flow with no additional stipulations.

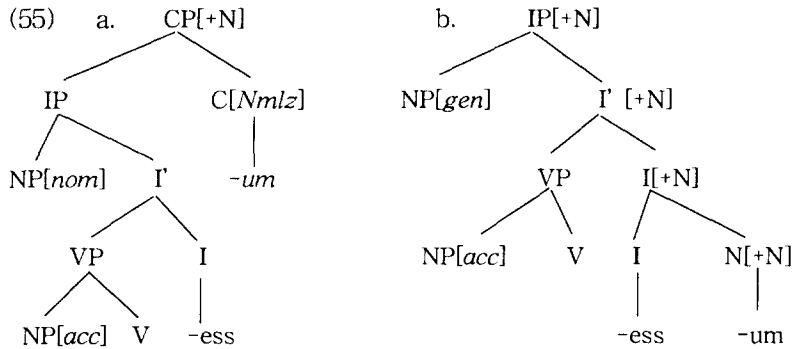
Here seems to be a proper place to compare our approach

⁶In fact, the only constraint needed to give this sort is the change of the HEAD value. The constraint on the SUBCAT and CONT(ENT) value can be assigned to a supersort (e.g., *stem*) since objects of this sort all observe the same constraint.

with a syntactic view assumed in Yoon (1991). He observed that the GEN(itive) case can be assigned to the subject of the sentential complement, as can be seen in (54).

- (54) John-uy/i chayk-ul ilk-um-i nollapta.
 John-GEN/NOM book-ACC read-Nmlz-NOM surprising
 'John's reading books is surprising.'

To account for such case alternation possibilities, Yoon posits different structures for sentential nominalizations, depending on the case value of the subject of the nominalized sentence, as given in (55).



His claim is that since in (55b), the subject is in the SPEC of I' which carries [+N] feature, it gets GEN Case, while the subject in (55a) gets Nominative case, since I' in this structure is not [+N] marked. Even though such a syntactic structure may be able to assign correct Case to the subject, it suffers from several problems.

First, it is unclear how his analysis can assign syntactic structures for *-ki* nominalized sentences. They cannot have the

same syntactic structures as in (55), since *-ki* cannot occur with an I (tense) element in his terms. One possible way is to assume that the *-ki* takes a tenseless VP and becomes an NP. Then, this structural assumption places the subject of the nominalized sentence into the SPEC of the nominalized NP. Then, the subject NP structurally positioned in the SPEC of NP will be assigned GEN case. However, there seem to be no *-ki* nominalized sentences whose subject can be GEN case marked, as seen from the ungrammaticality of (56).

- (56) *Tom-un [ku-uy ttena-ki]-lul kitali-ess-ta.
 Tom-TOP he-GEN leave-Nmlz-ACC wait-Past-Decl

Thus, his multiple structures, originally posited to capture case alternations, are not enough to capture such a fact. But our lexical approach can attribute the property of case alternations to the purely lexical property. Unlike the *non-factive* sort, the sort *factive* has its own specification such that the subject complement that its STEM selects for can be lexically marked GEN as well as NOM.

One further specific problem can be found in (55b) where *-um* is assumed to be an element of I, and the projection of which forms an IP. If this is the case, the question is why there can be no further CP projection in such constructions, i.e., why no C suffixes can be added further, as seen in (57).

- (57) *cap-usi-ess-um-ko 'catch-Hon-Past-Nmlz-COMP4'

The syntactic approach needs to assume a negative specification saying that no complementizers can select for IP

projection with [+N] marked. But in our lexical view, it is automatically borne out: a nominalizer suffix is attached to a verb stem, and the resulting stem is a noun. Thus, the lexical view naturally predicts the impossibility of further verbal suffixation after sentential nominalizers.

4. Conclusion

In sum, I hope to have shown that we can maintain the lexical integrity principle once we construct a suitable model of lexical structure. We have seen that the constraint-based theory of grammar, HPSG, which employs the mechanism of a 'sort hierarchy' constructed on the basis of well-defined *sorts*, can provide ways of enriching lexical information and obviated the need for syntax to refer to the internal structure of the word. We have also checked the plausibility of our analysis with several mismatch phenomena in Korean such as tense distribution, COMP selection and nominalization. Though it is true that these phenomena have strongly motivated the syntactic view, the results of the proposed lexical analysis in analyzing these phenomena show that without violating the principle of lexical integrity the lexical view can account for them at least as well as or better than the syntactic approach.

References

- Ahn, Hee-Don and Hang-Jin Yoon 1989. "Functional Categories in Korean" in S. Kuno *et al.*, eds., *Harvard Studies in Korean Linguistics* 3, 79-88.

- Anderson, Stephen. 1992. *A-Morphous Morphology*. Cambridge: Cambridge University Press.
- Bresnan, Joan and Sam A. Mchombo. 1995. "The Lexical Integrity Principle: Evidence from Bantu," *Natural Language and Linguistic Theory* 13, 181-254.
- Cho, Young-mee and Peter Sells. 1995. "A Lexical Account of Phrasal Suffixes in Korean," *Journal of East Asian Linguistics* 4, 119-74.
- Chomsky, Noam. 1991. "Some Notes on Economy of Derivation and Representation," in R. Friedin ed., *Principles and Parameters in Comparative Grammar*, 417-54.
- Han, Eunjoo. 1991. *Honorification in Korean*. Stanford University manuscript.
- Jung, Yeun-Jin. 1991. "Functional Projection, Specification and Parameters," in S. Kuno *et al.*, eds., *Harvard Studies in Korean Linguistics* 4, 281-94.
- Kathol, Andreas. 1994. *On Agreement in HPSG*. The Ohio State University manuscript.
- Lapointe, Steven. 1990. "Two Analyses of Korean Verb Inflections." in Y. No and M. Libucha eds., *Proceedings of ESCOL* 7, 187-203.
- Park, Byung-Soo. 1988. "How to Handle Korean Verbal Suffixes: A Unification Grammar Approach," in *Papers from the Sixth International Conference on Korean Linguistics*. ICKL, 593-600.
- Pollock, Jean-Yves. 1989. "Verb Movement, Universal Grammar, and the Structure of IP," *Linguistic Inquiry* 20, 365-424.
- Pollard, Carl and Ivan Sag. 1987. *Information-Based Syntax and Semantics*. Stanford, Calif.: CSLI, Stanford University.
- Pollard, Carl and Ivan Sag. 1994. *Head-Driven Phrase Structure*

- Grammar*. The University of Chicago Press.
- Sells, Peter. 1995. "Japanese and Korean Morphology from a Lexical Perspective," *Linguistic Inquiry* 26, 277-325
- Yoon, James. 1991. *Korean Nominalizations and Morphosyntactic Interaction*. University of Illinois, Urbana-Champaign manuscript.
- Yoon, James. 1993. *Tense, Coordination, and the Clausal Structure of English and Korean*. University of Illinois, Urbana-Champaign manuscript.
- Yoon, James. 1995. "Nominal, Verbal, and Cross-Categorical Affixation in Korean," *Journal of East Asian Linguistics* 4, 325-56.
- Yoon, James and Jeongme Yoon. 1990. "Morphosyntactic Mismatches and the Function-Content Distinction," in K. Deaton, M. Noske, and M. Ziolkowski eds., *Papers from the 26th Regional Meeting*. Chicago, Chicago Linguistic Society.

Department of English Language and Literature
Kyung Hee University
1 Hoegidong, Dongdaemoon-gu,
Seoul 130-701
E-mail: jongbok@nms.kyunghee.ac.kr
TEL: +82-2-961-0892

received: May 5. 1998.

accepted: June 29. 1998.