Mixed Categories and Multiple Inheritance Hierarchies in English and Korean Gerundive Phrases*

Chan Chung, Jong-Bok Kim, Byung-Soo Park, and Peter Sells

December 13, 2001

Abstract

One of the main puzzles in English and Korean gerundive phrases (VGPs) is that they display a mix of nominal and verbal properties. This has provided a challenge to syntactic analyses with a strict version of X-bar theory. Various approaches (cf. Lapointe 1993, Yoon 1996a & b, Kaiser 1998, Malouf 2000) have been proposed to solve this puzzle, but they all have ended up abandoning or modifying fundamental theory-neutral principles such as endocentricity, lexicalism, and null licensing (cf. Pullum 1991) with the introduction of otherwise unmotivated abstract elements or structures. This paper has shown that it is possible to analyze English and Korean VGPs in a way that maintains the lexical integrity principle (no syntactic rule refers to word-internal structure), captures endocentricity (the generalization that every phrase has a head), and avoids empty categories. This has been achieved through the framework of HPSG. HPSG is a sign-based grammar in which the basic unit of linguistic object *sign* is a structured complexes of linguistic information, represented by *typed feature structure*. The grammar of a language is based on the interactions of declarative constraints on types of signs. In capturing linguistic generalizations in a precise and concise manner, linguistic types are arranged into a multiple-inheritance hierarchy. The mechanism of multiple inheritance hierarchy allows a succinct way of encoding generalizations

*This paper was initiated from Kim (2001). We are grateful to Hee-Rhak Chae, Incheol Choi, Sae-youn Cho, Hyonoo Lee, Yongkyoon No, Byung-Soo Park, Eun-Jung Yoo for criticisms and helpful discussions. We also thank two anonymous reviewers of this journal for their criticisms and suggestions. All errors remain ours. This work was supported by Korea Research Foundation Grant (KRF2000-042-A00003)

about English and Korean gerundive constructions. This system provides a clean, streamlined way of capturing the mixed properties of English and Korean gerundive phrases. It also has allowed to have a simpler grammar that can capture parametric differences between the two languages.

1 Mixed Properties of English and Korean Verbal Gerundive Phrases

One of the main puzzles in English verbal gerundive phrases (VGPs) is that they display a mix of nominal and verbal properties. With regard to nominal properties, they can occur in syntactic positions that generally only admit NPs. For example, they can appear as the complement of a preposition as in (1)a, as a clause-internal subject as in (1)b/c, and as the focus of a cleft as in (1)d (cf. Pullum 1991, Malouf 2000, inter alia):

(1) a. They didn't approve of [my leaving without a word].
b. Tom believes that [John's taking a leave of absence] bothers Mary.
c. Why does [John's taking a leave of absence] bother Mary?
d. It's [John's taking a leave of absence] that bothers Mary.

However, the internal syntax of VGPs exhibit verbal properties, too: they can take accusative NP complements (like the verbs they are derived from), can be modified by adverbial modifiers, and can be negated with the negator *not* (cf. Pullum 1991, Malouf 2000, inter alia):

(2) a. Tom's calling (*of) the roll started each day.
b. Tom disapproved of my quietly/*quiet leaving before anyone noticed.
c. Tom's not having bathed for a week disturbed the other diners.

Like English VGPs, Korean VGPs also exhibit verbal properties internally and nominal properties externally. It is not difficult to find out that they exhibit verbal properties in terms of the internal syntax. One telling piece of evidence comes from the inheritance of arguments from the lexeme verb from which the gerundive verb is derived. As shown in (3), the gerundive verb takes the same arguments, a nominative subject and an accusative object:

(3) [John-i ecey ku chayk-ul/*uy
John-NOM yesterday that book-ACC/*GEN

ilk-ess-um]-i myenghwak-hata
read-PAST-NMLZ-NOM clear-do
'John's having read the book yesterday is clear'

Various other phenomena also show that VGPs are internally similar to VPs. The VGP can include a sentential adverb as in (4)a; an adverbial element can modify the gerundive verb as in (4)b; the phrase can include the sentential negation marker *an* as in (4)c; it also can contain the full range of auxiliaries as in (4)d, the phrase allows scrambling of its elements as in (4)e:

(4) a. John-i **papokathi** ku chayk-ul ilk-ess-um
John-NOM foolish that book-ACC read-PAST-Nmlz
'John's having read the book foolish'

b. John-i chayk-ul **ppalli/*ppalun** ilk-um
John-NOM book-ACC fast(adv)/*fast(adj) read-Nmlz
'John's reading books fast.'

c. John-i chayk-ul **an** ilk-um
John-NOM book-ACC NEG read-Nmlz
'John's not reading books.'

d. John-i chayk-ul ilk-ko **siph**-um
John-NOM book-ACC read-COMP want-Nmlz
'John's wanting to read books'

e. **ku chayk-ul** John-i __ ilk-ess-um(-i nollapta)
book-ACC John-NOM __ read-PST-NM-NOM surprising
'It is surprising that John read the book.'

Whereas the internal syntax of the VGPs is much like that of VPs, its external syntax is more like that of NPs. VGPs can appear in the canonical NP positions such as subject or object as in (5)a or as a postpositional object in (5)b (cf. Yoon 1996, Kaiser 1998)

(5) a. [ai-ka chayk-ul ilk-um]-i nollapta
child-NOM book-ACC read-NMLZ-NOM surprising
'That child's reading a book is surprising'

b. [John-i enehak-ul kongpwuha-m]-**eytayhay** mollassta
John-NOM linguistics-ACC study-Nmlz-about not.know
'(We) didn't know about John's studying linguistics.'

One thing worth pointing out here is that the VGP does not have the full distribution of NPs. As demonstrated in (6), the VGP cannot serve as the head of a relative clause, implying that the external syntax of the VGP is somewhat different from that of a canonical NP.

(6) *John-un [[salam-tul-i __ molulila-ko sayngkakha-n]
John-TOP people-PL not.know-COMP think-REL

[Mary-ka ilccik ttenass-um]]-ul alassta.
Mary-NOM early left-NMLZ knew
'*John knew [Mary's leaving early] that he thought that people wouldn't notice'.

Extraction is another instance indicating that the VGP behaves more like Ss and less like NPs in terms of the external syntax. Unlike a canonical nominal construction in (8), an element of the VGP can be extracted out as shown in (7).

(7) **ku chayk-ul** na-nun [John-i __ ilkess-um]-ul mitnunta
that book-ACC I-TOP John-NOM __ read-NMLZ-ACC believe
'That book, I believe John read.'

(8) **ku chayk-un* [na-nun [[Tom-i __ cwu-n] salam-ul] manassta]
the book-TOP I-TOP Tom-NOM give-PNE person-ACC met
'*The book, I met the person to whom Tom gave __ .'

These mixed and complicated properties of VGP (verb gerundive phrases) both in English and Korean have provided a challenge to syntactic analyses with a strict version of X-bar theory. Various approaches (e.g., Pullum 1991, Lapointe 1993, Malouf 2000, Yoon 1996, Kaiser 1998, among others) have been proposed to solve this puzzle, but they all have ended up abandoning or modifying fundamental theoretical conditions such as endocentricity, lexicalism, and null licensing (cf. Pullum 1991); see section 2.1.1 for elaboration.

This paper attempts to provide a new constraint-based HPSG approach that allows the conceptually natural considerations just mentioned to be maintained. Following Malouf (2000), we also assume that the traditional notion of taking grammatical categories as primitives is not enough, in particular for mixed categories like gerund, and that the mechanism of multiple classificational of category types with systematic inheritance can provide an effective and systematic way of capturing the mixed and complicated properties we find in both English and Korean VGPs. The analysis also provides

parametric differences between English and Korean VGPs on one hand and constructional differences among subtypes of the VGPs found in the two languages.

2 Morphosyntactic Properties of the Nominalizers

2.1 English

2.1.1 Basic Properties and Methodological Considerations

It seems to be natural to suppose that the *-ing* in an English VGP is a suffix. As noted by Pullum (1991) and others, the *ing* suffix in an English VGP, unlike the *-ing* in lexical nominalizations, cannot be claimed to be a nominalizer. One clear piece of evidence concerns that the V-*ing* form does not admit the nominal plural suffix *-s*:[1]

(9) a. the (frequent) singing-s of the aria (lexical)
b. *John's singing-s the aria (VGP)

Also, the VGP does not occur with the nonverbal negative *no* while admitting the verbal negative *not*.

(10) a. no/*not recording of the aria (lexical)
b. John's *no/not recording the aria (VGP)

Also as noted by Chomsky (1970), unlike the the lexical nominalization as shown in (11), the gerundive V-*ing* in verb-particle constructions behaves just like a verb as in (12):

(11) a. Chris's writing up of the paper
b. *Chris's writing of the paper up

(12) a. Chris's hastily writing the paper up
b. Chris's hastily writing up the paper

These morphosyntactic properties show us that the *-ing* in the VGP is a verbal suffix rather than a nominal suffix.

[1] As an anonymous reviewer points out, a semantic condition might block the attachment of the plural suffix here. What the data in (9) show is that the form *singing* can have the plural suffix when used as a true nominal element, whereas the same form cannot have the plural suffix in gerundive constructions.

As we have seen, the V-*ing* behaves like a verb and internally exhibits verbal syntax, whereas the whole VGP shows the external characteristics of a noun phrase. An ideal analysis would be of course to capture these mixed verbal/nominal properties with no ad hoc mechanisms. Pullum (1991) proposes three main 'theoretical desiderata' that any analysis of English verbal gerunds should satisfy: strong lexicalism, endocentricity, and null licensing.

- Strong lexicalism: Syntactic operations do not have access to the internal structure of words.
- Endocentricity: Every constituent has (at least) one distinguished head daughter identified as its head.
- Null licensing: No phonologically zero constituent should be posited that is neither semantically contentful nor syntactically bound.

Though these desiderata could be viewed as theory-dependent, we take the position that a preferred analysis is one that meets these three principles (cf. Malouf 2000, Kaiser 1998).

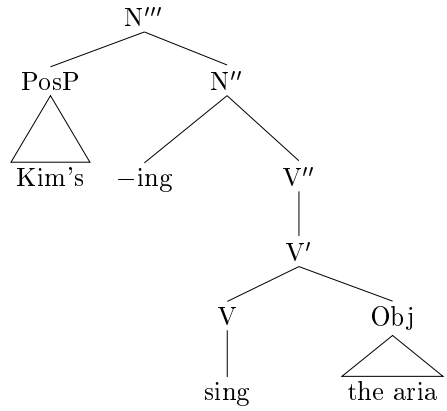
2.1.2 Previous Approaches

One earlier analysis for English VGPs is given by Jackendoff (1977). Observing that gerunds are difficult to fit into a standard X-bar syntax, he proposed the rule schema in (13).

(13) a. $X^i \rightarrow$ Affix V^i
b. $N'' \rightarrow$ -ing V''

These rules would then generate the structure (14) for a VGP like *Kim's singing the aria*:

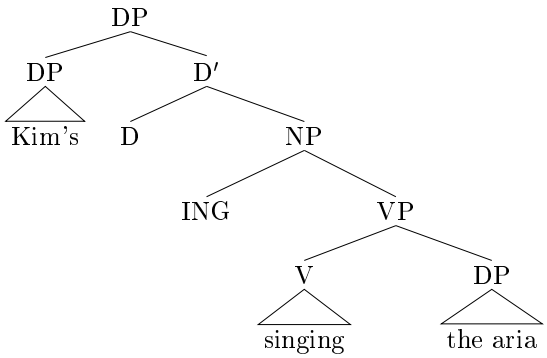
(14)



Lowering the affix *-ing* to the verb will generate the final output, *Kim's singing the aira.* This derivation, however, clearly violates the two main principles: lexical integrity and endocentricity of NP. There further seems to be no clear evidence to treat *-ing* as an independent syntactic element.

Avoiding the problem of lexical integrity, Abney (1987) introduces the abstract nominalizing morpheme *ING* to the *-ing* form of a verb and then assigns the following structure to a VGP like *Kim's singing the aria.*

(15)



Within this system, the abstract morpheme ING combines with the VP, resulting in a nonfunctional maximal projection NP. Since the subject is outside the nominalized part of the VGP, it takes a genitive phrase.

One immediate question that arises from such an analysis is what the head of the NP is: the NP has just ING and VP as its daughters. The introduction of the null phrasal head ING combining with the whole phrase also does not observe the principle of null licensing.[2]

Unlike such derivational approaches, Pullum (1991) proposes an analysis of English VGP within the framework of GPSG with the following rule:

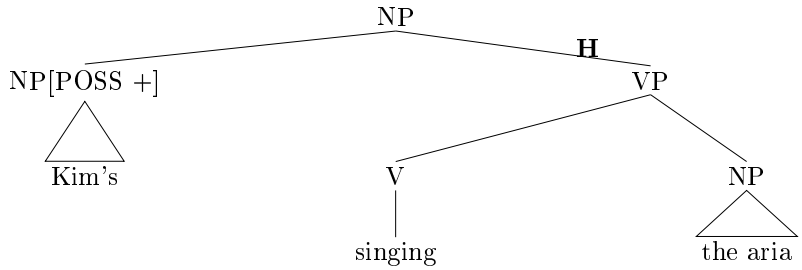
(16) N[BAR: 2] → (N[BAR: 2, POSS: +],) H[VFORM: prp]

Here the [VFORM: prp] identifies the present participle, the form of the verb paradigm that has the *-ing* suffix. The feature VFORM is tied to the category of verbs. What the rule basically specifies is that the element with the VFORM value is the head of the phrase, generating a structure like the following:[3]

[2] As noted by Malouf (2000), if English has a null nominalizing clitic or affix one would expect to find a language with richer morphology that has an overt nominalizing clitic. There exist few such languages. Yoon 1996b and Kaiser 1998 effectively propose that Korean nominalizations involve a kind of clitic, but see section 2.2 for arguments against such analyses.

[3] See Pullum (1991) for a detailed analysis.

(17)

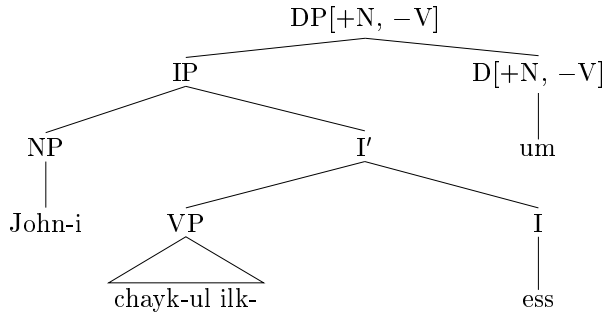


Though this structure avoids any phantom element in syntax while reflecting the description of VGPs as 'verbal inside and nominal outside', the rule (16) raises doubt as to the notion of headness.[4] Within this system, headness is achieved by allowing the VGP to dominate the genitive subject NP and the VP while designating this VP as the phrasal head. That is, although the mother NP and the head VP share no categorial properties, the VP can be designated as the phrasal head. Such a rule, being basically exocentric, violates one of the theoretical desiderata, the principle of endocentricity.

2.2 Korean

In analyzing the complicated and mixed properties of Korean nominalizations we have seen earlier, the first thing we need to look into is the properties of the nominalizers. There have been two main analyses of the nominalizers: as clitics (Kaiser 1998), and as phrasal affixes (Yoon 1989, 1996).[5] The structure in (18) is a canonical structure that a syntactic analysis (e.g., Yoon 1996) generates.[6]

(18)



[4]See Malouf (2000) for detailed discussion of this analysis.

[5]Strictly speaking, Kaiser's (1998) analysis claims that the nominalizer clitic is attached not to a sentence but to a verb lexeme in syntax.

[6]Yoon (1996b) adopts the following Phrasal Conversion Rule:

(i) *um/ki*: CAT: D (or N)
SUBCAT: $[V^n \_\_]^n$ where n is 0 or 2.

The approach in which the nominalizers are attached to phrases in syntax is motivated by syntactic factors such as the phrasal distribution of these morphemes, their productivity, and their scope (see section 4). For example, if one assumes that these morphemes are attached to syntactically-formed phrases rather than to roots or words, their phrasal or sentential 'scope' falls out naturally.

However, such syntactic and semantic behaviour need not necessarily lead us to the conclusion that the nominalizers are post-syntactic elements or clitics, in particular when considering their robust inflectional and lexical properties. The nominalizers are phonologically dependent and cannot occur in isolation. Neither can the verb stem occur alone.

Some of the basic properties of clitics (Zwicky and Pullum 1983) show that the nominalizers are not clitics. For example, the degree of selection between the clitics and the words preceding them is low. But the Korean nominalizers can attach only to untensed or tensed verb stems. They cannot be attached to any other syntactic category.

Another criterion of Zwicky and Pullum (1983) is that syntactic rules can affect affixed words but cannot affect clitic groups. But what we can observe in gerundive phrases is that the nominalized verb can undergo a gapping process as in (19). This entails that the verb stem forms a strong morphological unit with the attached nominalizer.

(19) [John-i sakwa-lul __] kuliko [Mary-ka panana-lul
John-NOM apple-ACC and Mary-NOM banana-ACC
mek-ess-um]
eat-PST-NMLZ
'John ate an apple and Mary a banana'

Another difficulty in treating nominalizers as clitics or as otherwise independent non-head elements comes from the fact that the meaning of the higher verb dictates the type of the nominalizer. In particular, the attachment of a delimiter such as *man* 'only' after the nominalizer as in (20) implies that the morphological form value of the nominalizer *-um/ki* should be visible to the higher verb.

(20) wuli-nun John-i nolayha-ki(/*um)-man-ul kitayhayessta
we-TOP John-NOM sing-NMLZ-DEL-ACC expected
'We expected John's singing only.

If the nominalizer were a clitic, an additional mechanism would need to be introduced to make the form value of this nominalizer available to the higher verb.

There is also no strong morphological evidence that the nominalizers are phrasal affixes. Lexical integrity tests, which show us the internal structure of words is opaque to various syntactic processes (Bresnan and Mchombo 1995 and Kim 2000), attest to the solid inflectional properties of these nominalizers.[7]

Extraction: According to the extraction test, no part of a word can be extracted. As in (21), the nominalized full word *hakca-taw-um* 'scholarliness' can be relativized but the *hakca* 'scholar' part of it.

(21) a. ku-ka __ wenha-n hakca-tawu-m
he-NOM want-REL scholar-Semi.Cop-NMLZ
'the scholar-likeness that he wanted'

b. *ku-ka __ -tawu-m-ul wenhayess-ten hakca
he-NOM __ -Semi.Cop-NMLZ-ACC wanted-REL scholar

Gapping: A verb can be gapped under identity with a verb in the previous sentence, but part of a verb cannot be gapped. As in (22), the verb must be gapped together with the nominalizer.

(22) *[John-i sakwa-lul __ -um] kuliko [Mary-ka panana-ul
John-NOM apple-ACC and Mary-NOM banana-ACC
mek-ess-um]
eat-PAST-NMLZ
'John ate an apple and Mary a bananna'

Coordination: No part of a verb can be factored out. No morphological constituents such as a nominalizer can function as a true conjunct.[8]

(23) a. [Tom-i nolay-lul ha-yess-um] kuliko
Tom-NOM sing-ACC do-PAST-NMLZ and
[John-i cwum-ul cwu-ess-um]
John-NOM dance-ACC dance-NMLZ
'Tom singing a song and John dancing'

b. *[Tom-i nolay-lul ha-yess-__] kuliko [John-i cwum-ul cwu-ess-um]

[7] The test for anaphoric islands appears to be irrelevant since the host is a verb stem.
[8] As in *[My mother and my father]'s house*, the clitic *'s* can be attached to a coordinated phrase.

If the nominalizer were simply taken to be an independent syntactic element that can be attached to a phrase like IP as in the syntactic analysis tree (**??**), (23)b would be an IP coordination.

The phenomena we have observed so far illustrate the strong inflectional properties of the nominalizers. Taking a syntactic or a clitic approach would require nontrivial modifications to capture these properties.

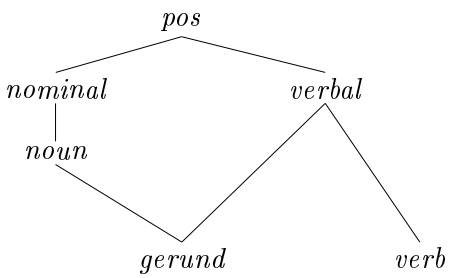
3 Analysis

In accounting for the properties of VGPs in English and Korean, we adopt the framework of HPSG, which factors syntactic properties into separate categorial, selectional, and constructional information. Within the system, the lexical head involves categorial information, and is projected into a phrase. This in turn means that categorial information will determine the external distribution of a phrase. Selectional information, encoded on the valence features of a lexical head, tells us what kind of other phrases the lexical head can appear with. Constructional information, encoded as constraints on particular constructions, determines the combination of syntactic constituents (see Malouf 2000, Sag and Ginzberg 2001).

3.1 English

As we have seen so far, the unusual properties of English VGPs are the combination of noun-like categorial properties with verb-like selectional properties. The starting point of our analysis is to posit the following part of speech hierarchy for English:

(24)



The hierarchy works as follows: in the hierarchy each phrase inherits constraints from its supertypes except the case that a default value is in conflict with a relevant non-default specification.[9] The hierarchy in (24) indicates

[9]For example, the Valence Principle can be overridden by a certain constraint such as the Head-Complement Schema. See Sag and Wasow 1999.

that verbs and gerunds are all subtypes of a common supertype called *verbal*. Features such as FORM, AUX and other features used in a variety of PSG analyses of English verbs are features defined as appropriate for the supertype *verbal* in the system (cf. Gazdar et al. (1982), Kim 2000).

One thing to note in the hierarchy is that the traditional head value is relevant to *noun* and *verb* only. Since *gerund* is a subtype of *noun* and *verbal*, but not *verb*, its head value is inherited from the supertype *noun*. The type *gerund* also has its own type-specific constraint on the FORM value. Thus, *gerund* will have the following consraints at least within the multiple inheritance hierarchy:

(25) $\begin{bmatrix} \textit{gerund} \\ \text{HEAD} \begin{bmatrix} \textit{noun} \\ \text{FORM gerund} \end{bmatrix} \end{bmatrix}$

Since *gerund* is now a subtype of *noun* with the HEAD value of *noun*, a phrase projected by a gerund will be an NP, occurring anywhere an NP is selected for. Thus, VGP will have the external distribution of NPs.[10]

The postulation of the FORM value has its own independent support in English. Specifically, the FORM value is crucial in selecting proper complements. For example, the verb *kept* selects a VP[prp] form:

(26) a. John kept [singing the song].
b. *John kept [sing the song].
c. *John kept [sung the song].
d. *John kept [sing the song].

There are also cases where a lexical element requires a complement with the value of [FORM gerund] (cf. Hudson 2000).

(27) a. It's/There is no use [telling him to leave/*a big fuss].
b. There is no point [telling him to leave/*anything else].
c. They prevented us from [finishing it/*its completion].
d. We had trouble [keeping the engine/*its function].

[10] There are some differences from Malouf (2000), which takes *gerund* as an independent syntactic category and gerunds are subtypes of *noun* and *relational*, not *verb*. The category *relational*, including *gerund*, *verb* and *adjective*, apparently has no direct reflex in syntax at all. Our analysis tries to avoid positing such an additional category that has a questionable syntactic status.

Capturing the external properties of VGPs, we now need to account for the internal properties repeated in (28):

(28) a. Verbal gerunds are modified by adverbs and not by adjectives.
b. A verbal gerund takes the same complements as the verb from which it is derived.

The present system requires no specification in the grammar for the properties in (28)a. The only thing we need to specify is that adverbs can modify elements of *verbal*. The examples in (29) explain this:

(29) a. The careful/*carefully restoration of the building took 10 years.
b. Tom's carefully/*careful restoring the building took 10 years.

The English verbal form is in general divided into finite (past and present) and nonfinite forms. The latter includes base, *-ing* form and (*-en* form) (for perfect and passive). In terms of the verbal form, we could observe that any verb with an *-ing* form, including the passive auxiliary *be* and the perfective auxiliary *have* will also have a verbal gerund form. Following this observation, we thus assume that any participle form of verb can be realized as the counterpart gerund form (either through lexeme realization or a lexical rule):

(30)
$$\begin{bmatrix} \textit{participle-verb} \\ \text{HEAD} \begin{bmatrix} \text{FORM prp} \end{bmatrix} \\ \text{SUBJ } \langle \boxed{1} \rangle \\ \text{COMPS } \boxed{b} \end{bmatrix} \Rightarrow \begin{bmatrix} \textit{gerundive} \\ \text{HEAD} \begin{bmatrix} \textit{noun} \\ \text{FORM gerund} \end{bmatrix} \\ \text{SUBJ } \langle (\boxed{1}) \rangle \\ \text{COMPS } \boxed{b} \end{bmatrix}$$

This rule produces as an output a lexical entry with the FORM value *gerund*. One difference from the input is that the subject is optional. Thus the system we set forth here naturally captures why the gerund inherits the complements of the verb it is derived from while making the subject optional. The data are repeated here:

(31) a. (Tom's) putting the book in the box appalled the teacher.
b. (Tom's) hammering the metal flat made a terrible noise.
c. Tom's discussing the issue so seriously surprised the teacher.

The other verbal properties of gerunds also follow naturally: since they are types of *verbal*, they cannot be pluralized or cooccur with particles; and they can be preceded by the sentential negator.

3.2 Korean

Given the morphological status of the nominalizers *-um* and *-ki* in section 2.2, the next question is what is the categorial status of the resulting morphological process. The attachment of a nominative or accusative marker to the gerundive verb may support its morphologically nominal status. However, a case marker can be attached to a verbal element too, as in (32).

(32) [John-i cip-ey ka-ss-nunka]-lul cosahaypoca
John-NOM home-LOC go-PST-Q-ACC investigate-SUG
'Let's find out whether John went home or not.'

There also exists negative evidence that casts doubts on the assumption that the VGP has the external syntax of canonical NP. As noted before, the VGP cannot be modified by a relative clause (see (6)) and some element can be extracted out of the phrase (see (7)). In addition, the phrase can neither serve as the host of genitive case nor attract plural marking, as shown in (33).[11]

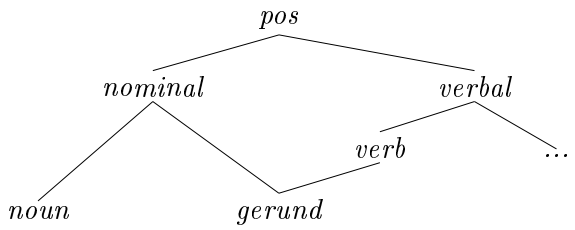
(33) a. *[John-i chayk-ul ilk-um]-uy cinsil
John-NOM book-ACC read-NMLZ-GEN truth

b. *[John-i chayk-ul ilk-um]-tul-i
John-NOM book-ACC read-NMLZ-PL-NOM

Such data suggest that though the gerundive verb may have some nominal properties, it is not a canonical noun.

To capture these mixed and complicated properties of Korean VGPs, we posit the following part of speech hierarchy, similar to the English one in (24):

(34)



[11] One could attribute this to the fact that the plural affix *tul* prefers a [ANIMATE +] host. But there are cases where it combines with a nonanimate noun as in *cinsil-tul* 'truth-PL'. See Kaiser 1998 for further discussion.

The classification in the hierarchy (34) assigns *gerund* as a subtype of *nominal* and *verb* with its own constraint on the FORM value. This means that *gerund* will have the following constraint minimally:

(35)
$$\begin{bmatrix} \textit{gerund} \\ \text{HEAD} \begin{bmatrix} \textit{verb} \\ \text{FORM gerund} \end{bmatrix} \end{bmatrix}$$

Since the type *nominal* is not specified with the part of speech value, *gerund* will inherit its part of speech value from *verb*.

This category classification can easily account for the mixed nominal and verbal properties of VGP constructions, too. The categorial properties of gerundive verbs are determined by their lexically specified head value. The present analysis defines the predicate with the nominalizer *-um/ki* in VGPs are all objects of *verb*. Under the mechanism of inheritance, a type will inherit all the constraints from its supertypes, and so this implies that the type *gerund* bears all the properties of its supertype *verb*. For example, the proposed system allows the lexeme *ilk-ta* 'read' can be realized as the gerundive verb *ilk-um* 'read-NMLZ' as in (36)b (either through a lexical rule or a lexeme realization constraint):

(36) a. ilk- 'read'

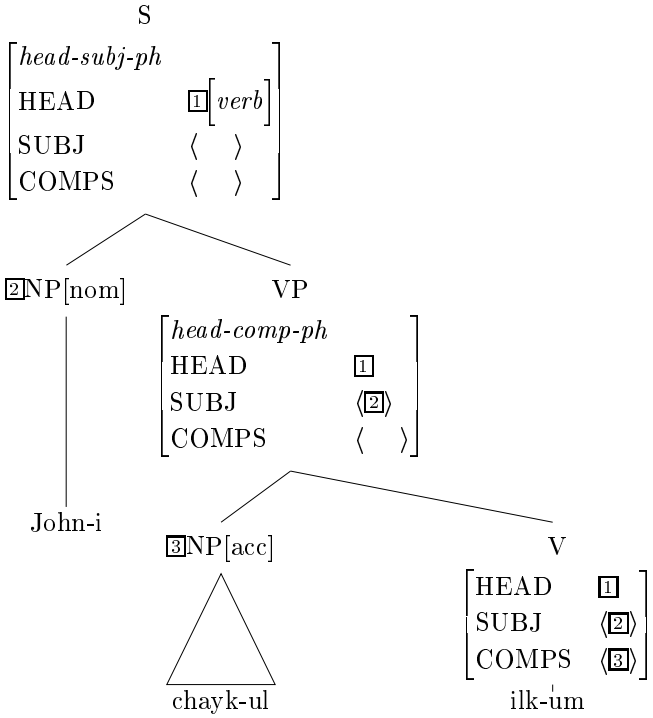
$$\begin{bmatrix} \textit{verb} \\ \text{FORM ilk-} \\ \text{SYN} \begin{bmatrix} \text{HEAD} \left[\textit{verb}\right] \\ \text{SUBJ } \langle \boxed{2}\text{NP} \rangle \\ \text{COMPS } \langle \boxed{3}\text{NP} \rangle \end{bmatrix} \end{bmatrix}$$

b. ilk-um 'read-NMLZ'

$$\begin{bmatrix} \textit{gerund} \\ \text{MORPH} \begin{bmatrix} \text{ROOT ilk-} \\ \text{I-FORM ilk-um} \end{bmatrix} \\ \text{SYN} \begin{bmatrix} \text{HEAD} \begin{bmatrix} \textit{verb} \\ \text{FORM gerund} \end{bmatrix} \\ \text{SUBJ } \langle \boxed{2}\text{NP} \rangle \\ \text{COMPS } \langle \boxed{3}\text{NP} \rangle \end{bmatrix} \end{bmatrix}$$

This process of lexical realization encodes a close relationship between *verb* and *gerund*. The present system with such a lexical specification projects a structure like (37) for VGPs.

(37)

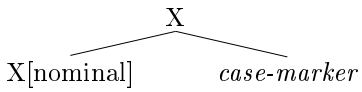


The gerundive verb differs syntactically from the lexeme in that its FORM value is added. This gerundive inherits all the other properties such as argument structure value from the lexeme. This explains why the gerund selects a nominative subject, can be modified by an adverb, allows sentential adverbials within the clause, combines with the sentential negative marker, occurs with an auxiliary verb, and the like (see section 1). Because the gerundive verb selects the same complement(s) as the verb lexeme it is derived from, the phrase formed by the gerundive and its complements will be a VP, forming a *head-comp-ph* (as can be seen in the VP structure of (37)). And since the gerund selects a subject, it will be eligible to head a *head-subj-ph*, which combines a head VP with a nominative subject. This is what the top node S in (37) represents, reflecting the internal properties of VGPs.

Various morphological and syntactic phenomena further support the line of our assumption. Support for assuming VGP predicates with the *um/ki* nominalizer as a subtype of *verb* rather than as a subtype of *noun* comes from (a) the presence of a tense and an agreement suffix and (b) the possibility of heading an independent sentence as in (38), which is one of the main differences from English VGPs.

(38) sensayngnim-i chayk-ul **ilk-usi-ess-um**.
teacher-NOM book-ACC read-HON-PAST-Nmlz
'The teacher has read the book.'

The external nominal properties of *gerund* come from its status as a subtype of *nominal*. We claim that case markers can combine with any elements belonging to the type *nominal*, as represented in the following structure (see (36)b):

(39)


Given the assumption that verbs with a nominalizer, a complementizer, or a question marker are all defined to be type of *nominal*, we can predict all these can be the complement of a case marker or a postposition as in (40).[12]

(40) a. John-i sakwa-lul **mek-ci-lul** moshayessta
John-NOM apple-ACC eat-COMP-ACC couldn't
'John couldn't eat apples.'

b. wuli-nun [John-i kacang cekhaphan-ka]-**eytayhay**
We-TOP John-NOM most appropriate-Q-about
nonuyhayessta
discussed
'We discussed whether John is the most suitable.'

The analysis further provides a simple way of capturing relativization and extraction phenomena. Though VGPs externally act like noun phrases, they do not allow a restrictive relative clause to modify them as repeated here in (41).

(41) *wuli-ka mall-in [John-i ku chayk-ul ilk]-um
we-NOM dissuade John-NOM that book-ACC read-Nmlz
'*John's reading the book that we dissuaded'

[12]This line of analysis also accounts for the fact that certain adverbials, which in the present analysis carry *nominal* value can cooccur with a case marker. See Wechsler and Lee 1996.

In the present system, the only thing we need to adopt is the independent constraint that a relative clause modifies a phrase projected from a *noun* element (not *nominal*). Then since a gerundive phrase is a projection of *verb* and *nominal*, we do not expect examples like (41). As noted before, it is possible to extract an element from VGPs, which is unexpected when considering the external status of the VGP to be a nominal phrase. But in our analysis, since the resulting VGP is a type of *head-subj-ph* we predict it to behave just like sentences (see the contrast in scrambling in head-subj-ph and in *noun-poss-ph* in section 4.2.) This prediction is borne out from the similarities between a gerundive clause (42)a and a true sentential complement clause (42)b:

(42) a. **ku chayk-ul** na-nun [John-i yelsimhi __
that book-ACC I-TOP [John-NOM hard
ilk-ess-um]-ul hwaksinhanta.
read-PST-Nmlz]-ACC convinced
'As for the book, I am convinced that John thoroughly read it.'

b. **ku chayk-ul** na-nun [John-i yelsimhi __
that book-ACC I-TOP [John-NOM hard __
ilk-ess-ta]-ko tulessta.
read-PST-DECL]-COMP heard
'As for that book, I heard that John read it hard'

The analysis therefore explains the external nominal properties of the construction we can find in Korean VGPs.

4 Constructional Differences

4.1 English

In English, there exist three subtypes of VGPs depending on the type of the subject: Genitive, Accusative, and *PRO*, respectively.

(43) a. I am proud of [Pat's winning the game].
b. I dislike [Pat painting her].
c. Pat is well known for [*PRO* painting flowers].

Of these three, GEN-VGPs and ACC-VGPs exhibit several observable differences that we need to account for. As noted by Abney (1987) and Malouf (2000), GEN-VGPs behave more like NPs while ACC-VGPs more like Ss.

One difference comes from agreement behavior in coordination:

(44) a. [That Pat came] and [that Chris left] bothers/?bother John.
b. [Pat] and [Chris] *bothers/bother John.

(45) a. [Pat's coming] and [Chris's leaving] ??bothers/bother John.
b. [Pat coming] and [Chris leaving] bothers/?? bother John.

Conjoined Ss tend to require a singular verb whereas conjoined NPs require a plural verb. The same pattern can be found in GEN-VGPs and ACC-VGPs, as shown in (45). In addition, these two types cannot be conjoined easily.

(46) a. *[Pat's coming] and [Chris leaving] bothers/bother me.
b. *[Pat coming] and [Chris's leaving] bothers/bother me.

Extraction also seems to display a difference between the two types. As in (47), it seems to be possible to extract a complement from an ACC-VGP but not from a GEN-VGP (see Horn 1975, Malouf 2000):

(47) a. Which city do you remember [him describing __]?
b. *Which city do you remember [his describing __]?

One more difference concerns pied piping as noted in Abney (1987) and Malouf (2000):

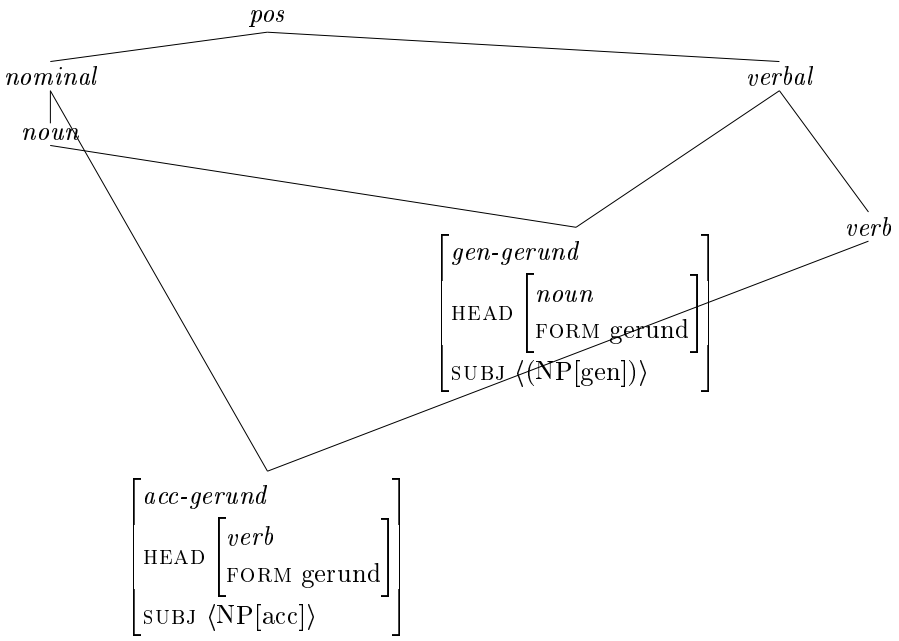
(48) a. I wonder [whose failing the exam] surprised the instructor.
b. *I wonder [who(m) failing the exam] surprised the instructor.

The same contrast can be found between NPs and Ss:

(49) a. This is the reporter [whose success] surprised John.
b. *This is the reporter for whom to win the Pulitzer Prize surprised Sandy.

These contrasts, extensively discussed in Malouf (2000), show that GEN-VGPs have something in common with NPs while ACC-VGPs are closer to Ss. To capture these contrasts, we extend the part-of-speech hierarchy in (24) to the one in (50):

(50)



What we can observe here is that *acc-gerund* and *gen-gerund* are identical with respect to the FORM value. But they are different in that the former is a subtype of *nominal* and *verb* whereas the latter is a subtype of *noun* and *verbal*. This classification eventually assigns the two types different HEAD values and case values to the subject: *acc-gerund* is *verb* with an accusative subject NP whereas *gen-gerund* is *noun* with an genitive subject NP.[13]

We claim that these different head values generate NP-like GEN-VGPs and S-like ACC-VGPs. For example, the participle verb form *painting* could be realized either as a *gen-gerund* or *acc-gerund* as in (51):

(51)

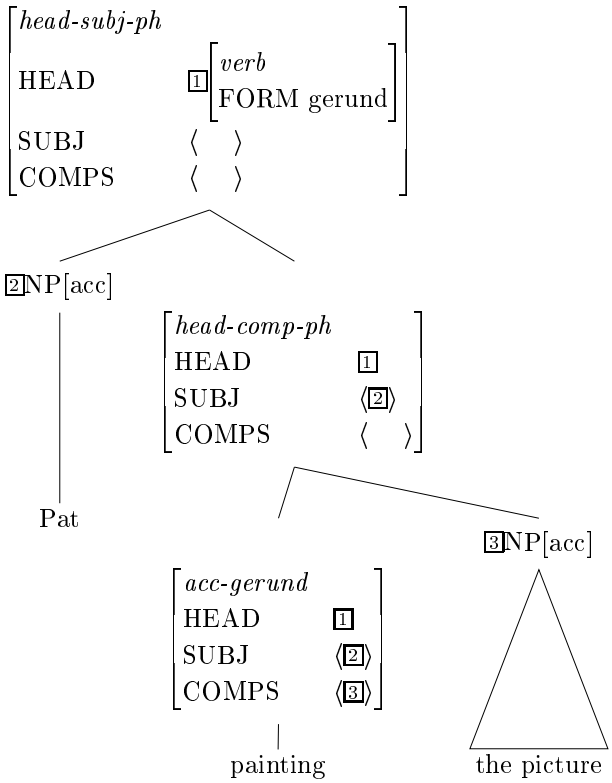
a.
$$\begin{bmatrix} \textit{gen-gerund} \\ \text{HEAD} \begin{bmatrix} \textit{noun} \\ \text{FORM gerund} \end{bmatrix} \\ \text{SUBJ } \langle(\text{NP[gen]})\rangle \\ \text{COMPS } \langle \text{NP} \rangle \end{bmatrix}$$

b.
$$\begin{bmatrix} \textit{acc-gerund} \\ \text{HEAD} \begin{bmatrix} \textit{verb} \\ \text{FORM gerund} \end{bmatrix} \\ \text{SUBJ } \langle(\text{NP[acc]})\rangle \\ \text{COMPS } \langle \text{NP} \rangle \end{bmatrix}$$

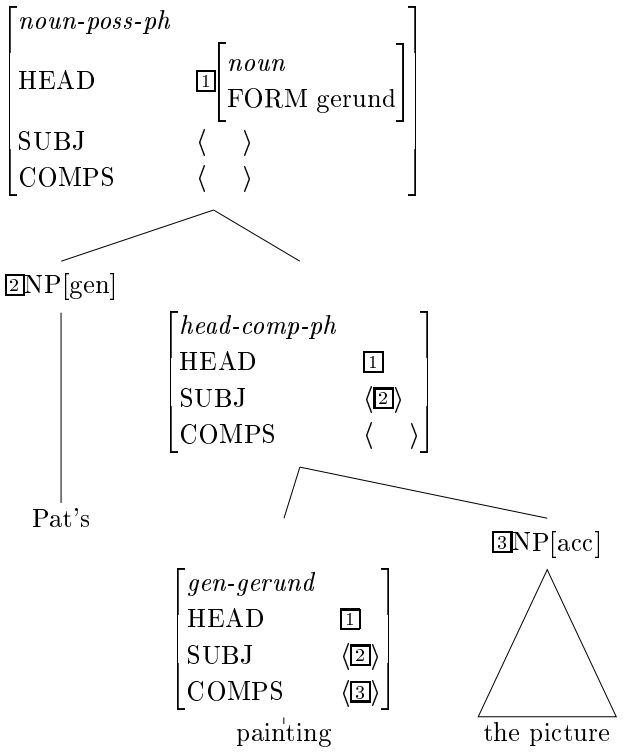
These lexical entries will project different structures as in (52) and (53).

[13] We assume that the case assignment to the elements with no lexically assigned case value as in gerundive constructions is done by general principles (or constraints) on the level of argument structure. See Przepiórkowski 1999 for a detailed analysis within this line.

(52)


head-subj-ph
HEAD
1
verb
FORM gerund
SUBJ
COMPS
2 NP[acc]
head-comp-ph
HEAD
1
SUBJ
2
COMPS
Pat
3 NP[acc]
acc-gerund
HEAD
1
SUBJ
2
COMPS
3
painting
the picture

(53)



Just by allowing enriched classifications for gerunds within the multiple hierarchy system, we could predict the differences between the two types of English VGPs with no additional mechanism. Since the head value of *gen-gerund* is *noun*, the VGPs projected from this gerundive type will behave like NPs whereas the head value of *acc-gerund* is *verb*, those projected from this head will be Ss.

4.2 Korean

The Korean VGP has its own language particular properties too. The language has two types of clausal nominalizer, *-um* and *-ki*, each of which behaves differently in several respects. For example, unlike the nominalizer *-um*, the host of *-ki* cannot usually be a tensed stem as shown in (54).[14]

[14] Another language particular property is that the nominalizer *um* preceded by a tense marker can be used as a kind of mood marker as in (i).

(54) John-i cip-ul ttena-(*ass)-ki-lul yaksokhayessta.
John-NOM house-ACC leave-PST-NMLZ-ACC promised
'John promised to leave home.'

There also exist semantic distinctions between the two types of nominalizers. For example, a gerundive verb with the *-um* nominalizer requires the higher verb to be a factive predicate such as *hwuhoyha-ta* 'regret', *myongpaykha-ta* 'evident', *pwunmyongha-ta* 'clear' etc, whereas one with the nominalizer *-ki* combines with a nonfactive predicate such as *kitayha-ta* 'expect', *kecelha-ta* 'reject', *myonglyongha-ta* 'order', *yaksokha-ta* 'promise', and so forth, as illustrated in the following contrast:

(55) a. [John-i cip-ul ttena-ss-um-i/*ki-ka] myengpaykhata.
John-NOM house-ACC leave-PST-NMLZ-NOM evident
'It is evident that John left home.'

b. Na-nun [John-i tolao-ki/*um]-lul kitayhanta.
I-TOP John-NOM come.back-NMLZ-ACC expect
'I expect John to come back.'

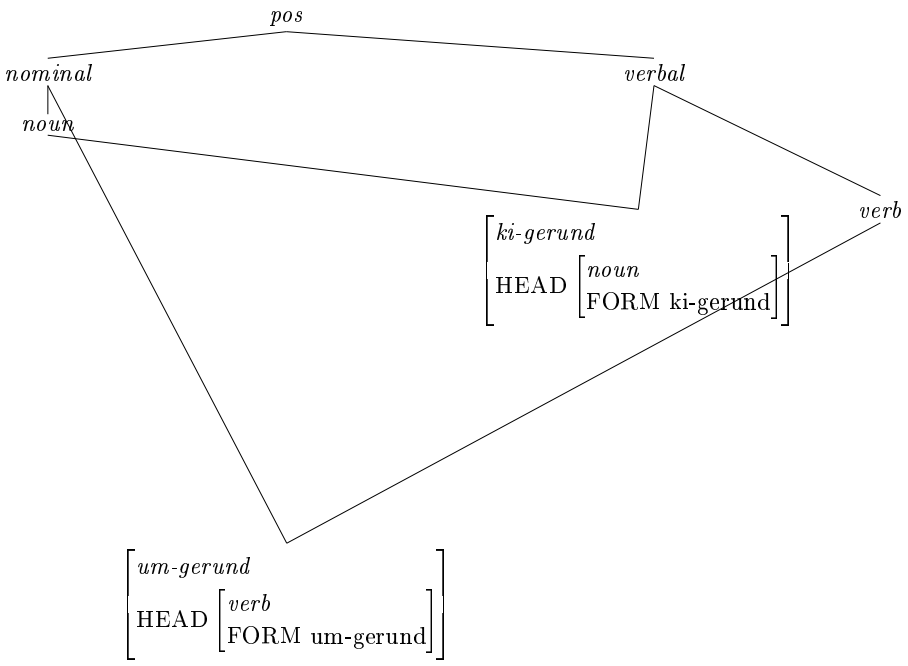
Further, as we have seen in (38), the *um* gerund can head an independent declarative sentence. This is not possible with the *ki* gerund:

(56) *sensayngnim-i chayk-ul **ilk-ki**.
teacher-NOM book-ACC read–Nmlz
'the teacher's reading the book'

As observed, one interesting constraint involving Korean VGPs is that the matrix predicate determines the type of nominalizer. As noted before, the restriction on the type of the VGP implies that the FORM value of gerundive predicates should be available to the higher predicate. One way to capture this is again to extend the Korean part-of-speech hierarchy in (34) as the one in (57):

(i) John-i ttena-ss-um.
John-NOM leave-PST-NMLZ
'John left.'

(57)



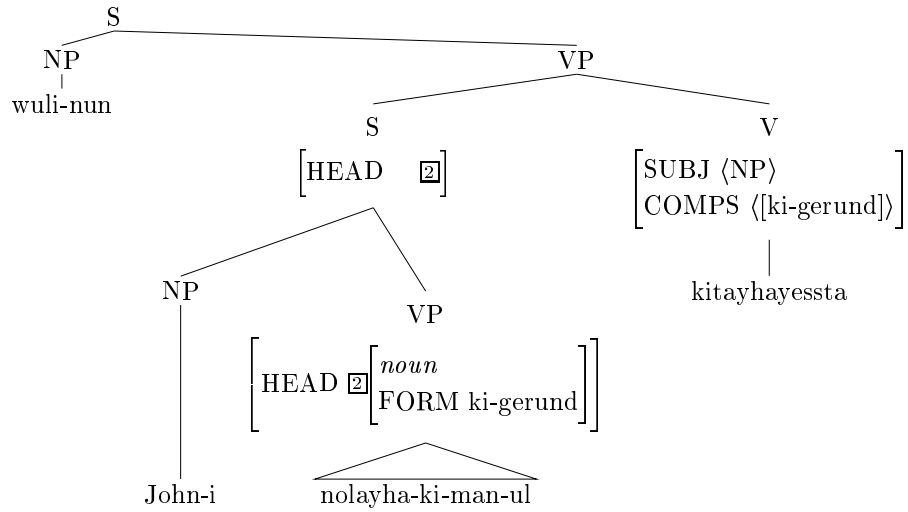
As the hierarchy indicates, we posit two different gerunds: *ki-gerund* and *um-gerund*. The former is *noun* with its own FORM value whereas the latter is *verb* with a different FORM value.

By so doing, we can capture the morphological selectional restriction in VGPs. An issue arises for examples like (55) where a delimiter occurs after the nominalizer. Either a clitic or a phrasal approach requires an additional mechanism to make the morphological form value of the nominalizer pass up to the matrix predicate. But the present analysis provides a straightforward account.

(58) wuli-nun John-i nolayha-ki(/*um)-man-ul kitayhayessta
we-TOP John-NOM sing-NMLZ-DEL-ACC expected
'We expected John's singing only.

The *ki-gerund* carries its own FORM value and this head feature projects up to the phrase, as presented in (59). This will eventually guarantee that the higher clause predicate can select a VGP whose predicate has the correct nominalizer value.

(59)



One phenomenon that appears to be a problem for lexical approaches, in that it seems to violate lexical integrity, involves examples like (60), in which the nominalizers seem to coordinate two sentences.

(60) [[John-i sakwa-lul mek-ess]-ko
John-NOM apple-ACC eat-PST-CONJ

[Mary-ka maykcwu-lul masi-ess]-m]
Mary-NOM beer-ACC drink-PST-NMLZ
'John ate apples and Mary drank beer.'

In our analysis, this is also predictable. Since the second VGP is also a type of sentence, cases like (60) are coordination of two sentences. There is no category mismatch in our analysis: the second conjunct is different from the first one only in its FORM value.

The analysis also provides a simple way of dealing with cases in which the subject is realized as genitive in both types:

(61) [John-uy chayk-ul ilk-um/ki]
John-GEN that book-ACC read-NMLZ-NOM
'John's reading books '

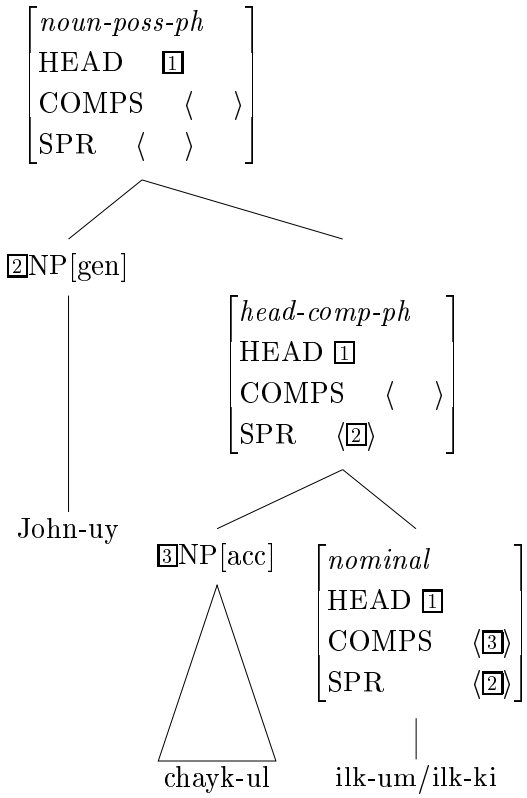
As we can observe, the case value on the subject of the gerundive verb is different from that on the subject of the English gerund. Korean allows only nominative or genitive. The example in (61) differs from the nominative subject VGP only in the way that the VP combines with a genitive specifier

to form a *noun-poss-ph*, whose constructional constraint can be represented as follow:

(62) *noun-poss-ph* → [CASE genitive], XP[*nominal*]

What this constraint means is that when the head is a type of *nominal*, its specifier can bear genitive case marking. Since both the *um-gerund* and the *ki-gerund* are subtypes of *nominal*, we can predict that their specifiers could carry genitive case. For example, this system will generate a structure like (63):

(63)



An independent constraint ensures that the combination of a head phrase with a specifier whose head is genitive is a type of *noun-poss(essive)-ph*. The head of this phrase is the gerund *ilk-um* or *ilk-ki*. It combines with its complement NP and forms a *head-comp-ph*, which then combines with a genitive specifier to form a *noun-poss-ph*.

This analysis allows us to predict similarities as well as contrasts between nominative subject VGPs and genitive subject VGPs. Since the head value of the resulting phrase in the latter is also *nominal*, we predict that it cannot serve as the head of a relative clause. This is what we observe as in (64).[15]

(64) *John-un [[salam-tul-i __ molulila-ko saynggakha-n]
John-TOP people-PL not.know-COMP think-REL

[Mary-uy ilccik ttena-m]]-ul alassta.
Mary-GEN early leave-NMLZ knew
'*John knew [Mary's leaving early] that he thought that people wouldn't notice'.

Since the genitive VGP is a type of *noun-poss-ph*, we expect no extraction as in a canonical *noun-ph*, and this is shown in (65). Unlike a *head-sub-ph*, we cannot extract an element out of a *noun-poss-ph*.[16]

(65) ***ku chayk-ul** na-nun [John-uy __ ilk-um]-i
that book-ACC I-TOP [John-GEN __ read-Nmlz]-NOM

nollapta
surprising
'(lit.) As for the book, John's reading is surprising.'

Another natural prediction would be the impossibility of coordinating the nominative subject VGP with a genitive type VGP. The impossibility of (66) stems from the coordination of a *head-subj-ph* and a *noun-poss-ph*:

(66) *[[John-i sakwa-lul mek-]-ko [Mary-uy maykcwu-lul
John-NOM apple-ACC eat-CONJ Mary-GEN beer-ACC

masi-]m]
drink-NMLZ
'(int)John eats apples and Mary's drinking beer'

[15]One potential advantage of our analysis is that it may predict speaker variation for examples like (64). For those who accept such examples, relative clause modification may depend on the type of construction, rather than on the head-value. That is, all nominal-constructions (including genitive VGPs) could be modified by a relative clause.

[16]For example, nothing prevents us from extracting an element from a *head-sub-ph* as in (i):

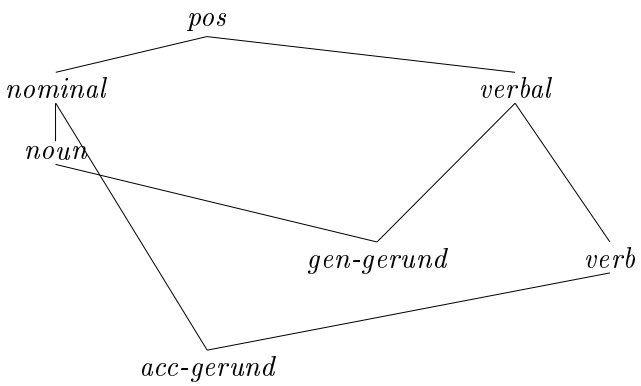
(i) **ku chayk-ul** na-nun [John-i __ ilk-ess-ta]-ko saynggakhanta
that book-ACC I-TOP [John-NOM __ read-PST-DECL]-COMP think
'That book, I think John read __ .'

As we have seen so far, once we have an articulated system for the two types of Korean gerund verbs within a multiple inheritance hierarchy, we can predict systematic differences between them as well as provide a clean analysis of various related phenomena.

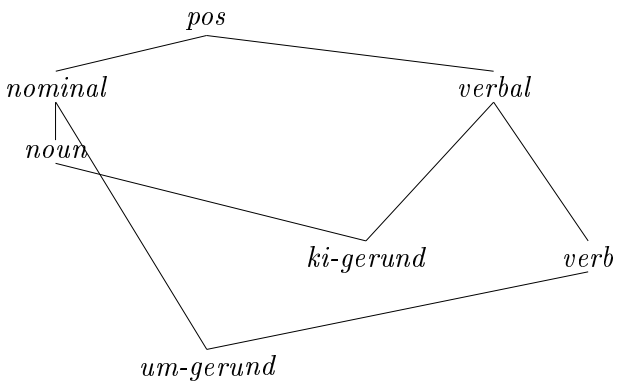
4.3 Parametric Differences between English and Korean

The treatment of the VGP in this sense implies that seemingly different VGPs in the two languages exhibit parametric similarities and differences. Both languages have at least two types of gerund: one bears more nominal properties whereas the other carries more verbal properties. This has been captured by an enriched hierarchy system as summarized in (67) and (68):

(67)



(68)



As indicated, we could predict that English *acc-gerund* is similar to Korean *um-gerund* whereas *gen-gerund* is similar to *ki-gerund*. We have also observed that language particular constraints and type-specific constraints on each type account for differences as well as idiosyncrasies.

5 Conclusion

This paper has shown that it is possible to analyze English and Korean VGPs in a way that maintains the lexical integrity principle (no syntactic rule refers to word-internal structure), captures endocentricity (the generalization that every phrase has a head), and avoids empty categories.

This has been achieved through the framework of HPSG. HPSG is a sign-based grammar in which the basic unit of linguistic object *sign* is a structured complexes of linguistic information, represented by *typed feature structure*. The grammar of a language is based on the interactions of declarative constraints on types of signs. In capturing linguistic generalizations in a precise and concise manner, linguistic types are arranged into a multiple-inheritance hierarchy. The mechanism of multiple inheritance hierarchy allows a succinct way of encoding generalizations about lexemes and phrases, while eliminating unnecessary stipulations. The enriched classification of grammatical categories with the inheritance system allowed us to build up a system where the principles of morphology are independent from those of syntax. The headness constraints on specific types (e.g., assigning HEAD value *noun* and *verb* but not on *nominal* or *verbal*) made it possible to have default inheritance (unlike Malouf's (2000) system), eventually bringing us the principle of endocentricity. The present system we have developed here requires no phonologically null elements whatsoever. The paper has further provided a clean, streamlined way of capturing the mixed properties of English and Korean VGPs. It also has allowed to have a simpler grammar that can capture parametric differences between the two languages.

References

Abney, Steven P. 1987. *The English Noun Phrase in Its Sentential Aspect.* Doctoral Dissertation, MIT.

Bresnan, Joan and Sam A. Mchombo. 1995. The Lexical Integrity Principle: Evidence from Bantu. *Natural Language and Linguistic Theory*, 13.2: 181–254.

Carpenter, Bob. 1992. Categorial Grammars, Lexical Rules, and the English Predicative. In R. Levine (ed.), *Formal Grammar: Theory and Implementation.* New York and Oxford: Oxford University Press.

Choi, Incheol. 1998. A Lexicalist Approach to Korean Sentential Nominalizations. Ms. University of Texas at Austin.

Chomsky, Noam. 1970. Remarks on Nominalization. In R. Jacobs and P. Rosenbaum (eds.), *Readings in English Transformational Grammar.* The Hague: Mouton. 184–221.

Gazdar, Gerald, Geoffrey K. Pullum, and Ivan Sag. 1982. Auxiliaries and Related Phenomena in a Restrictive Theory of Grammar. *Language* 58: 591–638.

Hudson, Richard. 2000. Gerunds and Multiple Default Inheritance. Ms.

Jackendoff, Ray. 1977. *X' Syntax: A Study of Phrase Structure.* Cambridge: MIT Press.

Kaiser, Lizanne. 1998. *The Morphosyntax of Clausal Nominalizations.* Doctoral Dissertation. Yale University.

Kim, Jong-Bok. 1998. Interface between Morphology and Syntax. *Language and Information* 2:1, 177-213.

Kim, Jong-Bok. 2000. *The Grammar of Negation: A Constraint-Based Approach.* Stanford: CSLI Publications.

Kim, Young-Joo. 1985. Semantic conditions for the Occurrence of Sentential Nominalizers Um and Ki. *Harvard Studies in Korean Linguistics I.* 168-177. Seoul: Hanshin.

Lapointe, Steven. 1993. Dual Lexical Categories and the Syntax of Mixed Category Phrases. In A. Kathol & M. Bernstein (eds.), *Proceedings of ESCOL*, 199-210.

Lapointe, Steve and Sarah Nielsen. 1994. A Reconsideration of Type III Gerunds in Korean. In *Japanese/Korean Linguistics V*, 305–320. Stanford: CSLI Publications.

Malouf, Rob. 2000. *Mixed Categories in the Hierarchical Lexicon.* CSLI Publications: Stanford.

Sag, Ivan. and Ginzberg, Jonathan. 2001. *Interrogative Investigations.* CSLI Publications: Stanford.

Przepiórkowski, Adam. 1999a. *Case Assignment and the Complement/Adjunct Dichotomy: A non-configurational constraint-based approach.* Ph.D. thesis, Universität Tübingen.

Pullum, Geoffrey. 1991. English Nominal Gerund Phrases as Noun Phrases with Verb-Phrase Heads. *Linguistics* 29: 763–799.

Wechsler, Steve and Yae-Sheik Lee. 1996. The Domain of Direct Case Assignment. *Natural Language and Linguistic Theory* 14.3, 629-664.

Yoon, James. 1989. *A Restrictive Theory of Morphosyntactic Interaction and Its Consequences.* Ph.D. Dissertation. University of Illiniois,

Urbana-Champaign.

Yoon, James. 1996a. A syntactic account of category-changing phrasal morphology: nominalizations in English and Korean. In H.D. Ahn et al. (eds.), *Morphosyntax in Generative Grammar*. Seoul: Hanshin, 63–86.

Yoon, James. 1996b. English Nominal Gerund Phrases as Phrasal Zero Derivations. *Linguistics* 34: 339–356.

Zwicky, Arnold M. and Geoffrey K. Pullum. 1983. Cliticization vs. Inflection: English n't. *Language* 59: 502–513.

Chan Chung
Department of English
Dongseo University
San 69-1 Julye-dong Sasang-gu
Pusan, Korea 617-716
cchung@dongseo.ac.kr

Jong-Bok Kim
School of English
Kyung Hee University
1 Hoegi-dong, Dongdaemoon-gu
Seoul, Korea 130-701
email: jongbok@khu.ac.kr

Byung-Soo Park
School of English
Kyung Hee University
1 Hoegi-dong, Dongdaemoon-gu
Seoul, Korea 130-701
email: byspark@khu.ac.kr

Peter Sells
Department of Linguistics
Stanford University
Stanford, CA 94305
email: sells@csli.stanford.edu