Two Types of Korean Light Verb Constructions in a Typed Feature Structure Grammar

Juwon Lee
Department of Linguistics, The University of Texas at Austin
juwonlee@utexas.edu

Abstract

In this paper, I present a lexical representation of the light verb *ha* 'do' used in two types of Korean *light verb constructions* (LVCs). These two types of the constructions have the typical theoretical and implementation problems as *multiword expressions* (MWEs): lexical proliferation of the possible light verb senses in the lexicon, potential overgeneration of ill-formed LVCs, and the semantic compositionality issue. Adopting and adapting the idea of qualia structure (Pustejovsky, 1991) into a typed-feature structure grammar (Copestake, 1993; Copestake, 2002; Sag et al., 2003), I suggest that some Korean common nouns have their associated predicate information in their lexical entries (e.g., the predicate meaning *cook* is included in the lexical entry of the common noun *pap* 'rice'). Thus such common nouns provide an appropriate predicate meaning to the light verb. The lexical constraints on the light verb and common nouns, and relevant phrase structure rules allow me to capture the generalizations and idiosyncrasies regarding LVCs in a systematic way.

1 Two Types of LVCs

A particular type of Korean LVCs, exemplified in (1), has been much studied (Chae, 1996, 2002; Choi and Wechsler, 2001; Kim et al., 2004; Kim et al., 2007, inter alia, and similar Japanese examples in Miyagawa, 1989; Matsumoto, 1996; Yokota, 2005, among others):

(1)a. ku-ka [swuhak-ul **kongpwu-lul**]
he-Nom math-Acc study-Acc
ha-yess-ta.[1]
do-Pst-Dec
'He *studied* mathematics.'
b. ku-ka [Mary-wa **tayhwa-lul**] **ha-**yess-ta.
he-Nom Mary-with talk-Acc do-Pst-Dec
'He *talked* with Mary.'

In (1a), the light verb *ha-yess-ta* 'do-Pst-Dec' requires as its complement the *verbal noun* (VN) phrase, *swuhak-ul kongpwu-lul* 'math-Acc study-Acc', and thus the types of LVCs in (1) are called VN-LVC in this paper, but see different syntactic analyses in Choi and Wechsler, 2001; Kim et al., 2004. Although the light verbs are the syntactic heads of the VN-LVCs, the core meanings of the sentences come from the verbal nouns. The mixed properties of VN in VN-LVC (that is, a VN can assign verbal cases to its arguments, but at the same time it can be modified by an adjective) have attracted much research on VN-LVCs (Grimshaw and Mester, 1988 on Japanese; Cho and Sells, 1991; Manning, 1993; Choi and Wechsler, 2001; Kim et al., 2007, among others).

However, there are many other usages of the Korean light verb *ha* 'do', which are almost ignored in the literature. In this paper, I investigate the two frequently-used, but less-studied types of Korean LVCs.

In the first type of the LVCs, the light verb requires a phrase headed by a *common noun* (CN) as its object (so, it is named CN-LVC here):

(2)a. ku-ka **pap-ul** **ha**-yess-ta.
he-Nom rice-Acc do-Pst-Dec
'He *cooked*/**ate* the rice (result product).'
b. ku-ka **khephi-lul**/***mwul-ul** **ha**-yess-ta.
he-Nom coffee-Acc/water-Acc do-Pst-Dec
'He *brewed /drank* the coffee/*water.'

In (2), we can see that the meaning of the light verb is determined by the object as with the VN-LVCs in (1).[2] Almost every VN seems possible to

[1] Abbreviations: Nom = Nominative, Acc = Accusative, Pst = Past, Dec = Declarative, Pass = Passive, Que = Question, Comp = Complementizer, Top = Topicalization, Rel = Relative marker

[2] Similar examples in English (Pustejovsky, 1991):

Proceedings of the Workshop on Multiword Expressions: from Parsing and Generation to the Real World (MWE 2011), pages 40–48,
Portland, Oregon, USA, 23 June 2011.

appear as the object in a VN-LVC. However, not every common noun can be the object of a CN-LVC.

The questions that naturally arise are 1) how to represent the light verbs of the CN-LVCs in the lexicon, and 2) how to formally and efficiently describe the way the predicate meanings (e.g., *brew* and *drink*) are derived from the objects (e.g., *khephi-lul* 'coffee-Acc').

If we treat CN-LVCs as words-with-spaces, then they suffer from a lexical proliferation in describing all possible meanings of the light verb expressions (e.g., *do_drink_coffee*, *do_brew_coffee*, *do_drink_tea*, *do_brew_tea*, etc.) (see Sag et al., 2002). On the other hand, a fully compositional analysis would overgenerate (e.g licensing **mwul-ul ha-yess-ta* 'water-Acc do-Pst-Dec' in (2b)) and would not be able to explain the problem of the semantic compositionality (that is, exactly where and how does the predicate meaning of the light verb phrase in a CN-LVC come from?) (see Sag et al., 2002). These problems of the CN-LVCs are not properly treated yet.

English LVCs have almost the same problems as the Korean CN-LVCs: idiosyncrasies on which light verb combines with a given noun (Abeille, 1988) (e.g., *make a mistake, give a demo*). A fully compositional account, on the other hand, would be unable to block alternative light verb combinations (e.g., **give a mistake*, **make a demo*) (see Sag et al., 2002).

Moreover, in Korean *serial verb constructions* (SVCs) the situation gets more complicated:

(3)a. ku-ka **pap-ul** **hay** **mek**-ess-ta.
he-Nom rice-Acc do eat-Pst-Dec
'He *cooked* the rice and *ate* it.'
b. ku-ka **khephi-lul*mwul-ul**/ **hay**
he-Nom coffee-Acc/ water-Acc do
masi-ess-ta.
drink-Pst-Dec
'He *brewed*/**drank* the coffee and *drank* it.'

In (3), the specific meanings of the light verbs depend on the common noun objects, which is parallel with the CN-LVCs. The difference, however, is that there is more restriction on the appropriate choice from the associated predicate(s) for the determination of the light verb meaning: e.g., only *brew* (creation sense) is allowed in (3b). I return to this semantic restriction in Section 3. The type of the constructions in (3) is called *serial verb-light verb construction* (SV-LVC) in this paper.

SV-LVCs have the same problems as CN-LVCs, including lexical proliferation of every possible senses of the serial light verb expressions with the words-with-spaces approach, the potential overgeneration, and the question of semantic compositionality.

These issues of the Korean LVCs as MWEs are crucial problems in *natural language processing* (NLP) like the disambiguation problems (see Sag et al., 2002). The goal of this paper is to solve the problems and to present an efficient formal account for CN- and SV-LVCs that is suitable for applications to linguistically precise NLP.

2 Grammatical Properties of CN-LVCs

CN-LVCs are very productive: the light verb *ha-* 'do' can combine with many (but not all) different common nouns to constitute CN-LVCs. The basic semantic and syntactic properties of CN-LVCs are discussed below.

2.1 Semantic Constraints of CN-LVCs

As is already illustrated in (2), there are two kinds of idiosyncratic restrictions on CN-LVCs. The first one is about what common noun can appear as the object in a CN-LVC:

(4)a. ku-ka **pap-ul**/***khwukhi-lul** **ha**-yess-ta.
he-Nom rice-Acc/*cookie-Acc do-Pst-Dec
'He *cooked* the rice/(int.) *baked* the cookie.'
b. ku-ka **khemphwuthe-lul**/***kaysanki-lul**
he-Nom computer-Acc/*calculator-Acc
ha-yess-ta.
do-Pst-Dec
'He *used* the computer/*calculator.'

The examples in (4) show that only certain food products or machines can occur as the objects in the CN-LVCs. The loan word *khwukhi-lul* 'cookie-

i) Mary **finished** the cigarette.
ii) Mary **finished** her beer.
iii) John **finished** the book.

The exact meaning of the verb is determined by the object: *finish smoking* for i), *finish drinking* for ii) and *finish writing* for iii). The verb, however, has also its own meaning: *finishing* X. So, in this case, the verb seems to be an intermediate type between light and heavy verbs.

Acc' in (4a) is not allowed, but other loan words, such as *khephi-lul* 'coffee-Acc' in (2b) and *khemphwuthe-lul* 'computer-Acc' in (4b), are fine. There seems to be no natural semantic class of common nouns that can appear in CN-LVCs, which leads me to attribute the idiosyncratic property to the individual common nouns.

The second idiosyncratic property is about what predicate is associated with what common noun. For instance, in (4a) the CN-LVC has only one reading, 'He *cooked* the rice', not other interpretations like 'He *ate* the rice,' although 'cook' and 'eat' are (at least semantically and maybe also statistically) plausible candidates for the associated predicates of the common noun *pap* 'rice'. Lapata (2001) uses a large corpus to acquire the meanings of polysemous adjectives (e.g., *fast*). However, such corpus findings only tell us the possible interpretations, but not impossible interpretations.

It seems intuitive that common nouns have such information about their related predicates since without a specific predicate given, we can normally guess what predicate might come after a common noun object in an incomplete sentence (at least in Korean whose word order is SOV) (see similar combinatoric information related with Korean VN of VN-LVCs in Cho and Sells, 1991 and Japanese VN in Manning, 1993).

In short, only some common nouns have such information about certain related predicates. Pustejovsky (1991) refers to this kind of relation as *cospecification*: i.e. like verb can select for its argument type, an argument also can select its associated predicates. The associated predicate information is included in the qualia structure of a lexical item (Pustejovsky, 1991). Among the four basic roles in qualia structure, the telic role has values about purpose and function of the object (e.g., *read* for *novel*), and the agentive role has values on factors involved in the origin or "bringing about" of an object (e.g., *write* for *novel*).

Building on the qualia structure, I propose that Korean common nouns have dual semantic components, the first of which is the meaning of the common noun itself, and second of which is the qualia structure. Details of the semantic feature structures of such common nouns are introduced in Section 5.

2.2 Syntactic Constraints of CN-LVCs

The CN-LVCs allow internal adverb modification:

(5)a. ku-ka pap-ul **ppalli** **ha**-yess-ta.
he-Nom rice-Acc quickly do-Pst-Dec
'He *quickly cooked* the rice.'
b. ku-ka khemphwuthe-lul **ppalli** **ha**-yess-ta.
he-Nom computer-Acc quickly do-Pst-Dec
'He *quickly used* the computer.'

So, the CN-LVCs are like Syntactically-Flexible Expressions (see Sag et al., 2002). I treat the CN-LVCs as a normal transitive verb phrase construction (generated by the general head-complement phrase rule) in syntax.

Since the light verb *ha* 'do' is syntactically a transitive verb, the passive counterparts of the CN-LVCs are predicted to be generated. However, only (4a) allows its passive:

(6)a. ku-eyuyhay **pap-i** **toy**-ess-ta.
he-by rice-Nom do.Pass-Pst-Dec
'The *rice* (product, not raw material) *was cooked* by him.'
b. *ku-eyuyhay **khemphwuthe-ka**
he-by computer-Nom
toy-ess-ta.
do.Pass-Pst-Dec

The passive light verb *toy* has the *become* meaning (i.e. creation sense). The associated predicate of *pap* 'rice' is *cook* (an agentive role predicate). Thus in (6a) *toy* is compatible with *be cooked*, which is also a "bringing about" predicate, but in the passive form. However, *khemphuthe* 'computer' has as its associated predicate *use* (a telic role predicate) and its passive form *be used* is also a telic role predicate. So, the creation meaning of *toy* is not compatble with the common noun subject *khemphwuthe-ka* 'computer-Nom' in (6b).

In sum, CN-LVCs are basically transitive phrases, but they are constrained by the semantic relations between common nouns and the light verb.

3 Grammatical Properties of SV-LVCs

As CN-LVCs are highly productive, SV-LVCs are accordingly very productive. The two types of the LVCs have similar semantic and syntactic constraints. But SV-LVCs are more restricted.

3.1 Semantic Constraints of SV-LVCs

As noted in (3), there are lexical constraints on the meanings of SV-LVCs. Consider (7):

(7)a. ku-ka **pap-ul** **hay** **ponay**-ess-ta.
he-Nom rice-Acc do send-Pst-Dec
(lit.) 'He *cooked* the rice and *sent* it (to me).'
b. ku-ka **khephi-lul** **hay** **ponay**-ess-ta.
he-Nom coffee-Acc do send-Pst-Dec
(lit.) 'He *brew* the coffee and *sent* it (to me).'

Since the common noun *pap* 'rice' has only one associated predicate, *cook* as shown in (2a), (7a) has only one reading. Although *khephi* 'coffee' has two associated predicates, *drink* and *brew* as evidenced in (2b), (7b) also has only one interpretation with *brewed* (the reading that he drank the coffee and sent it somewhere is implausible). Here, two hypotheses on the interpretations are possible: 1) any associated predicate that is plausible and available is chosen for the V1 light verb meaning, or 2) the V1 light verb meaning must be a creation (that is, an agentive role predicate).

The second hypothesis predicts that if a common noun has only a telic role predicate whose meaning is plausible in an SV-LVC, then the SV-LVC must be ill-formed. This is confirmed below:

(8) *ku-ka **khemphwuthe-lul** **hay**
he-Nom computer-Acc do
ponay-ess-ta.
send-Pst-Dec

The common noun *khemphuthe* 'computer' has the associated predicate *use*. The meaning of the telic role is plausible before the *sending* relation. So, the ungrammaticality of (8) rejects the first hypothesis.

Thus I suggest that certain common nouns have certain associated predicates information, and then in an SV-LVC, an available predicate of *bringing about* meaning must be chosen as the meaning of the V1 light verb *hay* in the construction. If such a predicate is not available, then the SV-LVC is ill-formed. Also, I have already illustrated that the agentive role predicate of a common noun is required for the generation of the passive CN-LVCs like (6a). Then how about passive SV-LVCs? I discuss this question in the following section.

3.2 Syntactic Constraints of SV-LVCs

First, adverbs can modify the serial verbs in the SV-LVCs:

(9)a. ku-ka pap-ul **ppalli** **hay** **mek**-ess-ta.
he-Nom rice-Acc quickly do eat-Pst-Dec
'He *quickly cooked* the rice and *ate* it.'
b. ku-ka khephi-lul **ppalli** **hay**
he-Nom coffee-Acc quickly do
masi-ess-ta.
drink-Pst-Dec
'He *quickly brew* the coffee and *drank* it.'

SV-LVCs are also categorized into Syntactically-Flexible Expressions. However, unlike CN-LVCs, the serial verbs (e.g., *hay mek-ess-ta* 'do eat-Pst-Dec') are complex predicates that need a special phrase (like (23) in Section 5).

As predicted, a common noun must have an agentive role predicate to license a well-formed passive SV-LVC. In other words, only if an SV-LVC is allowed, its passive SV-LVC is licensed:

(10)a. pap-i/khephi-ka **toy-e**
rice-Nom/coffee-Nom do.Pass-Comp
ponay-e **ci**-ess-ta.
send-Comp Pass-Pst-Dec
(lit.) 'The rice *was cooked* and *sent* (to me).'
(lit.) 'The coffee *was brewed* and *sent*
(to me).'
b. *khemphwuthe-ka **toy-e**
computer-Nom do.Pass-Comp
ponay-e **ci**-ess-ta.
send-Comp Pass-Pst-Dec

Just like the passive CN-LVCs, the exact meaning of *toy* depends on the common noun subject.

So, SV-LVCs are complex predicate structures in syntax, but they are also constrained by the semantics of common nouns and the light verb.

4 Pragmatic Factors

If a rich context is given, some ill-formed LVCs can be saved:

(11)a. ku-ka ***chayk-ul** **ha**-yess-ta.
he-Nom book-Acc do-Pst-Dec
b. ku-ka saygil **senmwul-lo** **chayk-ul**
he-Nom birthday present-as book-Acc
ha-yess-ta.

do-Pst-Dec
'he *gave* a book as a birthday *present*.'

The telic role of *senmwul* 'present' is *give* and this telic role seems to be passed to the object *chayk-ul* 'book-Acc' in (11b).

The grammaticality depends on what sense of a word is used in the sentence:

(12)a. *ku-ka **haksayng**-ul **ha**-yess-ta.
he-Nom student-Acc do-Pst-Decl
b. nwu-ka **haksayng** **ha**-lay?
who-Nom student do-Que?
'Who told you to *be* a student?'
(from the Korean TV show, *Hot Brothers*)

The ill-formed CN-LVC in (12a) can be saved in a special context where *haksayng-ul* 'student-Acc' is interpreted as a student role of a play (then the telic role *play* for the light verb), or in a colloquial context like (12b). Being a student (or lawyer, teacher, doctor, etc.) means that the person *performs* (telic role) the tasks of the position.

The object of the light verb can be implicit:

(13) **ce** **ken** twu-ko kan-ta. ne **hay**.
that thing leave-and go-Dec. you do.
'Let me leave that thing for you. You *have* it.'
(from the Korean movie, *Hello Ghost*)

The common noun object *ce ken* 'that thing' of the light verb is dropped from the second sentence of (13). The associated predicate of the common noun object is linked to the light verb across the sentence boundary. The abandonment of the possession of *that thing* seems to enforce the light verb to have the meaning of *have*. Such verbs as *write*, *cook*, *build* are related with physical creations, but *buy*, *have*, *possess* are related with relational creations.

Leaving the detailed formal analysis of the pragmatic factors for future research, I focus on the representations of the semantic and syntactic constraints.

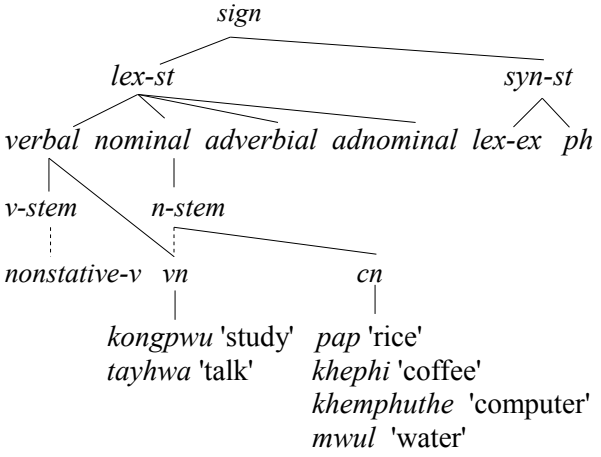
5 Typed-feature Structure Grammar

In this section, I present the formal analyses of the CN- and SV-LVCs in a typed-feature structure system (Copestake, 2002) based on the framework of the Head-driven Phrase Structure Grammar (Pollard and Sag, 1994; Sag et al., 2003).

5.1 Type Hierarchy of Korean

First, I adopt the following type hierarchy of the KPSG (Korean Phrase Structure Grammar) (Kim, 2004; Kim et al., 2004):

(14) Type hierarchy of linguistic expressions[3]:



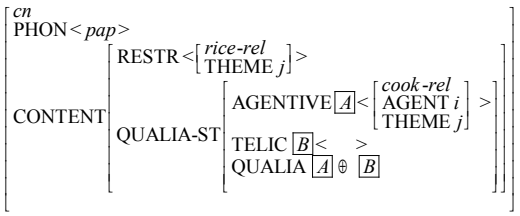
The type *vn* has the mixed properties inherited from its supertypes, *verbal* and *n-stem* (see Malouf, 1998, 2000; Choi and Wechsler, 2001). The type *cn* also inherits its constraints from its supertypes: for instance, nominal properties from the type *n-stem* (see Kim et al., 2004).

Briscoe et al. (1990) and Copestake (1993) illustrate some lexical entries with the qualia structure following Pustejovsky and Aniek (1988), Pustejovsky (1989, 1991). For example, *autobiography* has its associated predicates, *write* (the value of the agentive role) and *read* (the value of the telic role). They are represented in the lexical entry of *autobiography*.

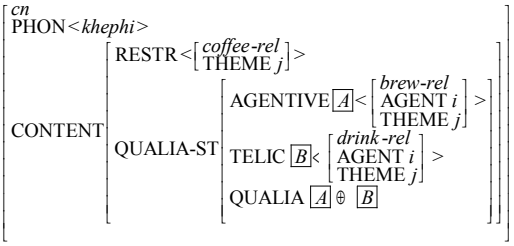
I declare that Korean common nouns have both the RESTR(ICTION) for normal semantics and the QUALIA-ST(RUCTURE), which in turn has the AGENTIVE and TELIC attributes, adopting the basic idea from Pustejovsky (1991) and adapting the feature structure from Copestake (1993). Moreover, I posit the QUALIA attribute whose value is the sum of the values of the AGENTIVE and TELIC. Based on this feature structure, I propose the following representations of the Korean common nouns:

[3] The dashed line here means that there are intermediate types between the types that are connected with it.

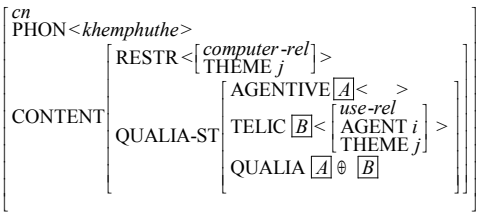
(15)a. Lexical entry for *pap* 'rice'



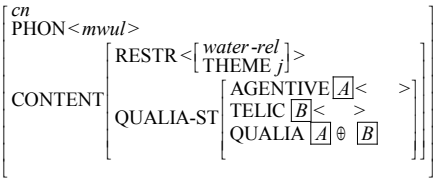
b. Lexical entry for *khephi* 'coffee'



c. Lexical entry for *khemphuthe* 'computer'



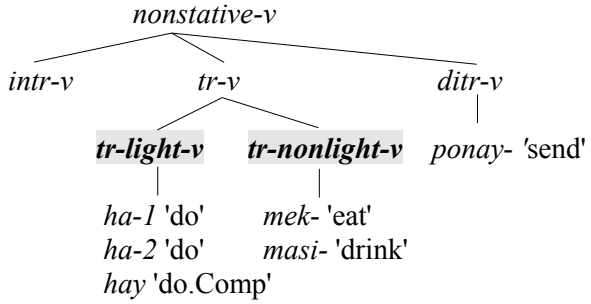
d. Lexical entry for *mwul* 'water'



In (15a), *pap* 'rice' has its associated predicate *cook* as the value of the AGENTIVE, but it has no value for the TELIC. Then, the QUALIA list must have only one value *cook*. In (15b), *khephi* 'coffee' has *brew* and *drink* in the AGENTIVE and TELIC, respectively. Then its QUALIA list includes *brew* as its first value, and *drink* as its second value. In (15c), the associated predicate of *khemphuthe* 'computer' is *use* (a telic role), which is then the sole value for the QUALIA. In (15d), *mwul* 'water' is declared not to have any value for the AGENTIVE or TELIC. Thus, it does not have a value for the QUALIA, either.

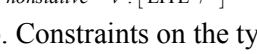
Now as for the relevant verbs of the LVCs, I divide the type *tr*(*ansitive*)-*v*(*erb*) in the following type hierarchy further into *tr*(*ansitive*)-*light-v*(*erb*) and *tr*(*ansitive*)-*nonlight-v*(*erb*):

(16) Type hierarchy of non-stative verbs:



Three lexical entries of the light verbs are under the type *tr-light-v*. They have different properties that can be captured by the following constraints:

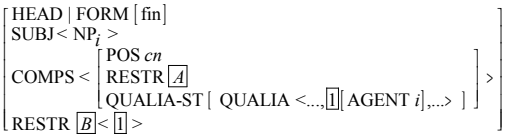
(17)a. Constraints on the type, *nonstative-v*:



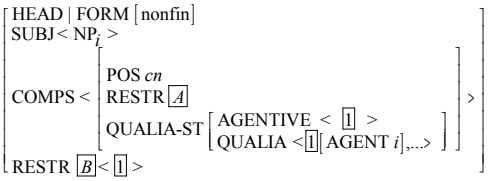
b. Constraints on the type, *tr-light-v*:

c. Constraints on *ha-1*:

d. Constraints on *ha-2*:



e. Constraints on *hay*:



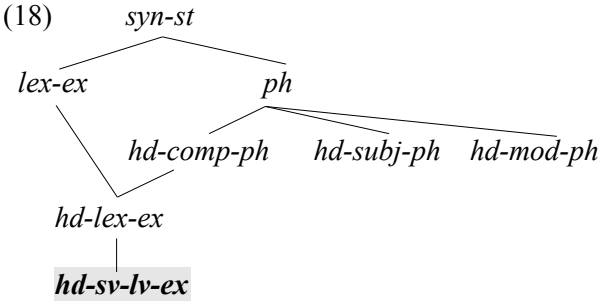
In (17a), the defeasible feature [LITE /–] is posited on *nonstative-v*. So, all the subtypes inherit [LITE /–], except for *tr-light-v* since in (17b), the defeasible feature [LITE /–] is overridden by the specification of the feature value. Only two types *tr-nonlight-v* and *ditr-v* can appear as V2 in SV-LVCs, and now they can be referred to as verbs that take at least one complement and have the feature [LITE /–]. In (17c), the RESTR of *ha-1* is claimed to be empty list since the light verb that combines with a verbal noun phrase does not seem to contribute a core meaning to the VP as shown in (1). However, in (17d), the meaning of *ha-2* is linked to a value of the QUALIA of the common noun object. This constraint of *ha-2* will guarantee that in CN-LVCs, any value in the QUALIA (e.g.,

drink or *brew* of *coffee*) can be chosen for the specific meaning of the light verb. Another effect of the constraint is preventing the common nouns like *mwul* 'water' from appearing in a CN-LVC since such common nouns are declared to not have a value for the QUALIA as in (15d). Finally, in (17e), a separate lexical entry for the V1 light verb *hay* is posited due to the different properties from *ha-2*: e.g., *ha-2* can get a tense, so is *finite* but *hay* cannot receive a tense, so is *nonfinite*. In addition, the meaning of the V1 light verb *hay* is identical only with the Agentive value of the common noun object.

5.2 Head-Complement Combinations

Along with the lexical entries, syntactic rules are needed. In the type hierarchy of (14), the relevant subtypes of *syn-st* are represented below (cf. Kim, 2004; Kim et al., 2004; Kim, 2010). I added the new type *hd-sv-lv-ex* as a subtype of *hd-lex-ex*:

(18)



The following general head-complement rule (see Sag et al., 2003; Kim 2004) generates a phrase of the type *hd-comp-ph*:

(19) Head-Complement Rule:

$$\mathrm{XP}[\textit{hd-comp-ph}] \rightarrow \boxed{1}, \mathbf{H}\left[\mathrm{COMPS} \langle \ldots, \boxed{1}, \ldots \rangle\right]$$

In addition to the syntactic head-complement phrase rule, the following semantic constraints on the structures are defined (Sag et al., 2003):

(20)Semantic Compositionality Principle:
In any well-formed phrase structure, the mother's RESTR value is the sum of the RESTR values of the daughters.

Equipped with the Head-Complement Rule and the Semantic Compositionality Principle, VPs in CN- and VN-LVCs can be generated:

(21)a. Head-Complement Phrase of CN-LVC:

$$\begin{bmatrix} \textit{hd-comp-ph} \\ \mathrm{COMPS} \langle\ \rangle \\ \mathrm{RESTR}\ \boxed{A} \oplus \boxed{B} \end{bmatrix} \rightarrow \boxed{1}, \mathbf{H}\begin{bmatrix} \textit{tr-light-v} \\ \mathrm{HEAD} \mid \mathrm{FORM}\ [\mathrm{fin}] \\ \mathrm{COMPS} \langle \boxed{1} \begin{bmatrix} \mathrm{POS}\ \textit{cn} \\ \mathrm{RESTR}\ \boxed{A} \\ \mathrm{QUALIA\text{-}ST}\ [\ \mathrm{QUALIA} \langle \ldots, \boxed{2}, \ldots \rangle\ ] \end{bmatrix} \rangle \\ \mathrm{RESTR}\ \boxed{B} \langle \boxed{2} \rangle \end{bmatrix}$$

b. Head-Complement Phrase of VN-LVC:

$$\begin{bmatrix} \textit{hd-comp-ph} \\ \mathrm{COMPS} \langle\ \rangle \\ \mathrm{RESTR}\ \boxed{A} \end{bmatrix} \rightarrow \boxed{1}, \mathbf{H}\begin{bmatrix} \textit{tr-light-v} \\ \mathrm{COMPS} \langle \boxed{1} \begin{bmatrix} \mathrm{POS}\ \textit{vn} \\ \mathrm{RESTR}\ \boxed{A} \end{bmatrix} \rangle \\ \mathrm{RESTR} \langle\ \rangle \end{bmatrix}$$

In (21), according to the Semantic Compositionality Principle, the VP in the CN- or VN-LVC has the sum of the RESTR values of the object and the light verb.

In the type hierarchy (18), the type *hd-comp-ph* has the subtype which is constrained by the following Head-Lex Rule (cf. Kim et al., 2004):

(22) Head-Lex Rule:

$$\begin{bmatrix} \textit{hd-lex-ex} \\ \mathrm{COMPS}\ \boxed{A} \oplus \boxed{B} \end{bmatrix} \rightarrow \boxed{1}\begin{bmatrix} \mathrm{LEX} + \\ \mathrm{COMPS}\ \boxed{A} \end{bmatrix}, \mathbf{H}\left[\ \mathrm{COMPS} \langle \boxed{1} \rangle \oplus \boxed{B}\ \right]$$

In (22), the head element combines with its complement, whose complements and some of head's complements are passed up to the resulting *hd-lex-ex*.

The constraints on *hd-lex-ex* are inherited to its subtype *hd-sv-lv-ex*. This phrase type is responsible for the combinations of the serial light verb expressions in SV-LVCs:

(23) Head-SV-LV-EX Rule:

$$\begin{bmatrix} \textit{hd-sv-lv-ex} \\ \mathrm{SUBJ} \langle \boxed{1} \rangle \\ \mathrm{COMPS}\ \boxed{B} \oplus \boxed{D} \\ \mathrm{RESTR}\ \boxed{C} \oplus \boxed{E} \end{bmatrix} \rightarrow \boxed{2}\begin{bmatrix} \textit{tr-light-v} \\ \mathrm{HEAD} \mid \mathrm{FORM}\ [\mathrm{nonfin}] \\ \mathrm{SUBJ} \langle \boxed{1} \rangle \\ \mathrm{COMPS}\ \boxed{B} \langle \boxed{4} \begin{bmatrix} \mathrm{RESTR}\ \boxed{A} \\ \mathrm{AGENTIVE} \langle \boxed{5} \rangle \\ \mathrm{QUALIA} \langle \boxed{5}, \ldots \rangle \end{bmatrix} \rangle \\ \mathrm{RESTR}\ \boxed{C} \langle \boxed{5} \rangle \end{bmatrix},$$

$$\mathbf{H}\begin{bmatrix} \textit{nonstative-v} \\ \mathrm{LITE}\ / - \\ \mathrm{SUBJ} \langle \boxed{1} \rangle \\ \mathrm{COMPS} \langle \boxed{2} \rangle \oplus \boxed{D} \\ \mathrm{RESTR}\ \boxed{E} \end{bmatrix}$$

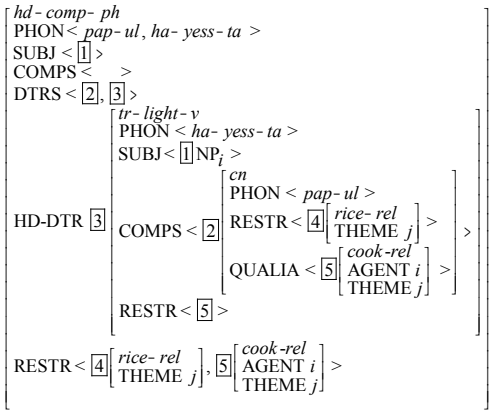
In (23), the nonstative verbs with [LITE /–] (which are not intransitive) like *eat*, *drink* and *send* require the V1 light verb *hay* as its complement.

Now, the serial light verb expressions (e.g., *hay mek-ess-ta* 'do.Pass eat-Pst-Dec') can be licensed with the Head-SV-LV-EX Rule and the Semantic Compositionality Principle. Furthermore, (23) can rule out the ill-formed SV-LVCs like (8) **ku-ka khemphwuthe-lul* [*hay ponay-ess-ta*] 'he-Nom

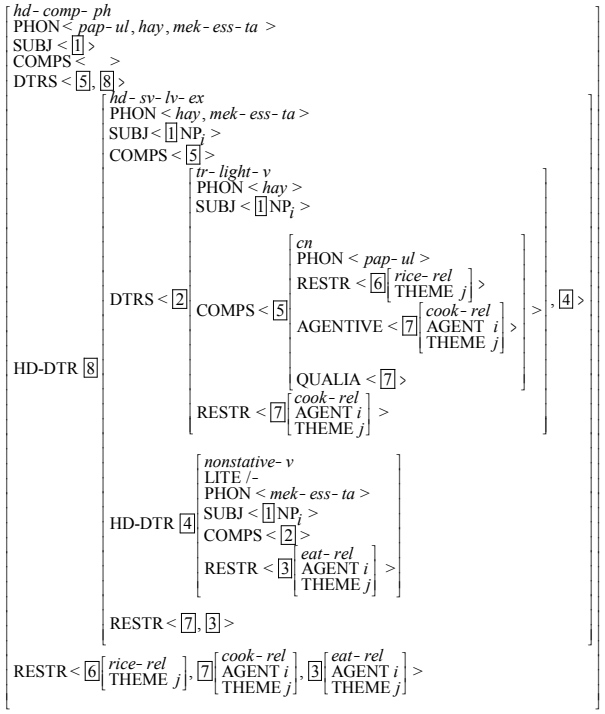
computer-Acc do send-Pst-Dec' because *hd-sv-lv-ex* requires a common noun object that has an AGENTIVE value, but *khemphwuthe* 'computer' has no value for it. The implausible interpretation '#He *drank* the coffee and sent it (to me)' for (7b) [*ku-ka khephi-lul* [*hay ponay-ess-ta*]] is also blocked since the meaning of the light verb *hay* is linked only to the AGENTIVE value of the object.

The following feature structures show the analyses of the VP [*pap-ul ha-yess-ta*] 'rice-Acc do-Pst-Dec' in the CN-LVC (2a) and the VP [*pap-ul* [*hay mek-ess-ta*]] 'rice-Acc do eat-Pst-Dec' in the SV-LVC (3a):

(24) [*pap-ul ha-yess-ta*]:



(25) [*pap-ul* [*hay mek-ess-ta*]]:



6 Conclusion and Future Work

The light verb *ha-2* 'do' is used for CN-LVCs and *hay* is used for SV-LVCs. I also proposed that certain Korean common nouns have their associated predicate meanings in the QUALIA-ST. These lexical constrains on individual common nouns and the light verbs, and the relevant phrase structure rules account for the regular and idiosyncratic properties of the two LVC constructions in a systematic manner.

I believe that the current analysis can possibly extend to the corresponding LVCs in other languages (especially Japanese since it has similar LVCs with the light verb *suru* 'do' and allows serial verbs). The VPs with the verbs *start* or *finish* (see Pustejovsky, 1991) can also be accounted for using the qualia structure: e.g., *pap-ul sicakhata*/ *kkuthnayta* 'start/ finish (**cooking**/*eating) the rice', *khephi-lul sicakhata*/ *kkuthnayta* 'start/ finish (**brewing**/*drinking) the coffee', *khemphuthe-lul sicakhata*/ *kkuthnayta* 'start/ finish (*building/**using**) the computer' and **mwul-ul sicakhata*/ *kkuthnayta*. My temporary hypothesis for such the VPs is that there is the ranking (that is, agentive role > telic role), so the agentive role of a common noun object is used first with *start* or *finish*, but if agentive role is not available, then telic role is used, and if even telic role is not available, then it is ungrammatical.

More comprehensive research with corpus data and the actual implementation of the analysis in the *Linguistic Knowledge Building* (LKB) system (Copestake, 2002) are left for future work.

Acknowledgment

I would like to thank Katrin Erk for her helpful comments. John Beavers and Stephen Wechsler have also greatly helped me to improve the paper. The insightful comments of the anonymous reviewers are gratefully acknowledged as well.

References

Anne Abeille. 1988. Light verb constructions and extraction out of NP in a tree adjoining grammar. In *Papers of the 24th Regional Meeting of the Chicago Linguistics Society*.

Ann Copestake. 1993. Defaults in Lexical Representation. In *Inheritance, Defaults, and the Lexicon*, 223-245. Cambridge: Cambridge University

Press.

Ann Copestake. 2002. *Implementing Typed Feature Structure Grammars*. Stanford, CA: CSLI Publications.

Briscoe Ted, Ann Copestake and Bran Boguraev. 1990. Enjoy the paper: Lexical semantics via lexicology. In *Proceedings of the 13th International Conference on Computational Linguistics*, 42-47.

Carl J. Pollard and Ivan A. Sag. 1994. *Head-Driven Phrase Structure Grammar*. Stanford, CA: CSLI Publications.

Christoper D. Manning. 1993. Analyzing the Verbal Noun: Internal and External Constraints. *Japanese/Korean Linguistics* 3, 236-253.

Hee-Rahk Chae. 1996. Light Verb Constructions and Structural Ambiguity. In *Proceedings of PACLIC 11: Language, Information and Computation,* 99-107.

Hee-Rahk Chae. 2002. Predicate Nominals and Light-er Verbs in Korean: An Indexed PSG Approach. In *Proceedings of the 2002 LSK International Summer Conference* I, 115-125.

Incheol Choi and Stephen Wechsler. 2001. Mixed Categories and Argument Transfer in the Korean Light Verb Construction. In *Proceedings of the 8th International Conference on Head-Driven Phrase Structure Grammar,* 103-120.

Ivan A. Sag, Timothy Baldwin, Francis Bond, Ann A. Copestake, and Dan Flickenger. 2002. Multiword expressions: A pain in the neck for nlp. In *Proceedings of the 3rd International Conference on Computational Linguistics and Intelligent Text Processing*, 1-15.

Ivan A. Sag, Thomas Washow, and Emily M. Bender. 2003. *Syntactic Theory: A Formal Introduction*, 2nd edition. Stanford, CA: CSLI Publications.

James Pustejovsky and Peter G. Aniek. 1988. On the Semantic Interpretation of Nominals. In *Proceedings of COLING-1988*, 518-523.

James Pustejovsky. 1989. Current Issues in Computational Lexical Semantics. In *Proceedings of 4th European ACL*, xvii-xxv.

James Pustejovsky. 1991. The Generative Lexicon. *Computational Linguistics* 17(4), 409-441.

Jane Grimshaw and Armin Mester. 1988. Light Verbs and Theta-Marking. *Linguistic Inquiry* 19, 205-232.

Jong-Bok Kim. 2004. *Korean Phrase Structure Grammar*. Seoul: Hankook Munhwasa.

Jong-Bok Kim, Jaehyung Yang, and Incheol Choi. 2004. Capturing and Parsing the Mixed Properties of Light Verb Constructions in a Typed Structure Grammar. In *Proceedings of PACLIC* 18, 81-92.

Jong-Bok Kim, Kyung-Sup Kim, and Jaehyung Yang. 2007. Structural Ambiguities in the Light Verb Constructions: Lexical Relatedness and Divergence. *The Linguistic Association of Korea Journal* 15(2), 207-231.

Jong-Bok Kim. 2010. Argument Composition in Korean Serial Verb Constructions. *Studies in Modern Grammar* 61, 1-24.

Kenji Yokota. 2005. The structure and meaning of Japanese light verbs. *Language Sciences* 27, 247-280.

Maria Lapata. 2001. A Corpus-based Account of Regular Polysemy: The Case of Context-sensitive Adjectives. In P*roceedings of the 2nd Meeting of the North American Chapter of the Association for Computational Linguistics*, 63-70.

Robert Malouf. 1998. *Mixed Categories in the Hierarchical Lexicon*, Linguistics Department, Stanford University, Ph.D. Dissertation.

Robert Malouf. 2000. *Mixed Categories in the Hierarchical Lexicon: Studies in constraint-based lexicalism*. Stanford, CA: CSLI Publications.

Shigeru Miyagawa. 1989. Light Verbs and the Ergative Hypothesis. *Linguistic Inquiry* 20, 659-88.

Yo Matsumoto. 1996. A Syntactic Account of Light Verb Phenomena in Japanese. *Journal of East Asian Linguistics* 5(2), 107-149.

Young-Mee Yu Cho and Peter Sells. 1991. A lexical account of phrasal suffixes in Korean. ms. Stanford University.