Copy constructions and their interaction with the copula in Korean

Jong-Bok Kim
Kyung Hee University

Peter Sells
Stanford University

Proceedings of the 12th International Conference on
Head-Driven Phrase Structure Grammar

Department of Informatics, University of Lisbon

Stefan Müller (Editor)

2005

Stanford, CA: CSLI Publications

pages 213–231

Kim, Jong-Bok & Peter Sells. 2005. Copy constructions and their interaction with the copula in Korean. In Stefan Müller (ed.), *Proceedings of the 12th International Conference on Head-Driven Phrase Structure Grammar, Department of Informatics, University of Lisbon*, 213–231. Stanford, CA: CSLI Publications.
DOI: 10.21248/hpsg.2005.12.

Abstract

We argue here for a lexicalist analysis of the Korean copula (following Kim et al. (2004))), on the basis of different properties of sequences of noun-plus-copula, which shows word-like behavior, in contrast to noun and negative copula, which are independent syntactic units. The interactions of these items with various copy constructions brings out their clear differences. The analysis is formalized in HPSG using Lexical Sharing, from Wescoat (2002).

1. The Copula

The Korean copula *-i-* forms a phonological word with its preceding N host (see e.g., Oh (1991), Cho and Sells (1995)); (1)a is a representative example. The negative copula *ani-* in (1)b shows a similar structure, but without the phonological cohesion; in fact its complement (*salam*) takes nominative case.

(1) a. ku haksayng-un ilpon-eyse o-n salam-i-ta
that student-TOP Japan-from come-PAST person-COP-DECL
'That student is a person from Japan.'

b. ku haksayng-un ilpon-eyse o-n salam-i ani-ta
that student-TOP Japan-from come-PAST person-NOM NCOP-DECL
'That student is not a person from Japan.'

The nominative marker has allomorphs *-i* and *-ka* which are regularly conditioned.

In canonical predicative uses the copula *-i-* does not allow case-marking on the complement N, nor any other final suffix such as *-(n)un* or *-to*, as shown in (2)a. The negative copula *ani-ta*, however, does allow such suffixes, as seen in (2)b.

(2) a. apeci-nun hakca(*-ka/*-nun/*-to)-i-ta
father-TOP scholar(*-NOM/*-TOP/*-FOC)-COP-DECL
'My father is a scholar.'

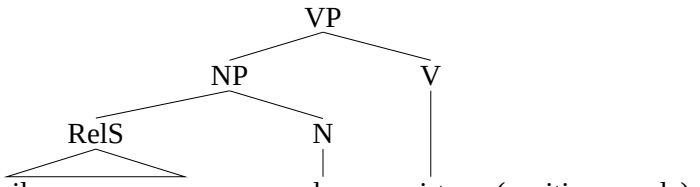
b. apeci-nun hakca(-ka/-nun/-to) ani-ta
father-TOP scholar(-NOM/-TOP/-FOC) NCOP-DECL
'My father is not a scholar.'

The impossibility of final suffixes seen in (2)a suggests the lexicality of the form consisting of the noun host plus copula (see Cho and Sells (1995), Sells (1997)), for it is not clear why a clitic treatment of the copula would predict the lack of parallel in the examples in (2).

[†]This paper represents a part of joint work with Michael T. Wescoat on the application of Lexical Sharing to Korean and Japanese. We are particularly grateful to Michael, and also to Ivan Sag, for discussions on the best formalization of Lexical Sharing within HPSG. Our paper has also benefitted from comments from the audience at HPSG05.

On the other hand, the modifier *ilpon-eyse o-n* in (1) forms a constituent with the head noun *salam*, showing evidence of an NP in the syntax. This observations seems most consistent with the view of the copula *-i-* as a clitic (the most thorough treatment is in Yoon (2003)). Specifically, the copula is treated as a V in syntax that forms a phonological word with its adjacent N host, the head of an NP complement, indicated by the '=' in (3). The negative copula has the same syntax but just happens not to be a clitic (hence the examples in (1) are syntactically parallel):

(3)



In this paper we add to a growing body of evidence which shows that noun plus copula is indeed a lexically-formed verb (see especially Kim et al. (2004)), despite the apparent evidence in favor of a clitic analysis. A fully lexical account can nevertheless allow that the N which hosts the copula can head a fully-formed syntactic NP (in (1)b), through the adoption of the Lexical Sharing approach of Wescoat (2002). Informally, Lexical Sharing allows words to instantiate *one or more* lexical-category nodes, and so, alongside familiar one-word-to-one-phrase instantiation, exemplified by *salam-i* 'person-NOM', the theory also posits *portmanteau words*, which instantiate two or more adjacent lexical-category nodes. This allows us to accept the lexicality of *salam-i-* 'person-COP', a form which may receive verbal inflectional affixation in the lexicon (see Kim et al. (2004)). We will adopt the syntactic structure for the positive copula in (3) while nevertheless treating the host noun plus copula as a single word.

Our evidence for lexicality involves the careful separation of several related 'copying' constructions in Korean, which provide evidence for lexical and syntactic units. In section 2 we briefly describe the first three of these constructions; then, in section 3, we introduce examples involving noun plus copula, and some of these involve a fourth construction. In section 4 we present the HPSG analysis which accounts for the data given in sections 2 and 3, with the exception of the specific analysis of the copula in terms of lexical sharing, which is given in section 5.

2. Evidence from the Echo Construction

Our new evidence regarding the status of the copula comes from the subtle contrasts that we can find between apparently similar examples involving copying various amounts of syntactic material. In all, we introduce 4 constructions in this paper, listed in (4); the remainder of this section is focussed on the first three types.

(4) Constructions introduced in this paper

a. Echo Contrastive Construction (ECC): V-*ki-nun* V-*ta*, sets up a negative implicature.

b. *Ha* Constrastive Construction (HCC): V-*ki-nun ha-ta*, sets up a negative implicature.

c. VP-Topic Construction: VP-*ki-nun* VP-*ta*, has no negative implicature.

d. Noun Copy Construction (NCC): N-*nun* N, indicates that N is a prototypical member of its class.

The first construction, 'Echo Contrastive Construction (ECC)', involves the doubling of Vs, but does not extend to their phrasal arguments or adjuncts.[1] As shown in the translation of (5)a, the ECC sets up a negative implicature in the interpretation of the whole sentence (see Choi (2003), Cho et al. (2004), Kim (2002), Aoyagi (2005)). This negative implicature is indicated by the '*but* . . . ' in our translations.

(5)b shows a related construction, the '*Ha* Contrastive Construction (HCC)', which involves using the verb *ha* ('do') for the second verb, rather than a copy of the first, but which also has the negative implicature.

(5) a. ECC: copy the verb root; inflect the second verb for tense etc.

John-i Tom-ul [manna-ki-nun manna-ss-ta]
John-NOM Tom-ACC meet-NMLZ-TOP meet-PAST-DECL
'John met Tom, but . . . .'

b. HCC: follow the verb root with a form of *ha*; inflect the second verb (*ha*) for tense etc.

John-i Tom-ul [manna-ki-nun hay-ss-ta]
John-NOM Tom-ACC meet-NMLZ-TOP do-PAST-DECL
'John met Tom, but . . . .'

The interaction of the ECC with the copula provides strong support for our claim about the lexicality of the copula. The only grammatical form of an ECC with the positive copula *-i-* also involves doubling the N host of the copula, as in (6)a (see Oh (1991), Kim and Chung (2002)). The copied parts are underlined.

[1] Strictly speaking, the ECC can copy a V^0 which may itself be internally complex, consisting of more than one word, but it may not contain phrasal material (see Cho et al. (2004); and (18) below).

The label 'predicate cleft' has been applied to the ECC, for example by Nishiyama and Cho (1998) and Jo (2004b). The term comes from Koopman (1984), who describes a construction in Vata which has a copy of the verb in initial position in the clause, followed by a full clause (which is SOV). We do not think that the Korean constructions that we discuss here have the same properties, either in terms of syntax or interpretation.

(6) a. ku salam-i [mikwuk-eyse kongpwu ha-n]
that person-NOM America-at study do-PAST
haksayng-i-ki-nun haksayng-i-ta
student-COP-NMLZ-TOP student-COP-DECL
'That person is a student who studied in America (but he still doesn't speak English well).'

b. *ku salam-i [mikwuk-eyse kongpwu ha-n]
that person-NOM America-at study do-PAST
haksayng-i-ki-nun i-ta
student-COP-NMLZ-TOP COP-DECL
(copying 'V' only; ungrammatical)

Copying the copula alone is completely impossible, as in (6)b. Now under the clitic analysis, the copula never forms a syntactic unit with its NP complement, and so we would expect (6)b to be the grammatical version. Additionally, it is unclearl how to make the copied part in (6)a a syntactic constituent – formally, it would correspond to the head of the complement NP and the following V (which selects for that NP). However, we see clearly that the ECC treats N+Copula as a lexical constituent, and that the copula alone cannot function as a pure V in the syntax, from the contrast in (6)a and (6)b.

The facts in (6) contrast directly with the ECC facts involving the negative copula *ani-*, which takes a nominative-marked complement (see (1)b): the *ani-* verbal part can be doubled by itself, as in (7)b, just like a regular verb (cf. (5)a). And while the doubling of N + negative copula as in (7)a is grammatical, this example does not have the 'negative implicature' interpretation typical of the ECC, but rather has a VP-topic interpretation – along the lines of 'as for not being a fool, that person is not a fool'.[2] This asymmetry shows that the ECC targets a verb in the syntax and intuitively copies it, meaning that there is a lexical form *haksayng-i-* for (6)a alongside *ani-* for (7)b. The positive copula is in fact one of a class of verbal elements including *-tap-* 'is every bit' and *-kath-* 'seem' (noted by Yoon (2003)), which behave in the same way, including in the ECC.

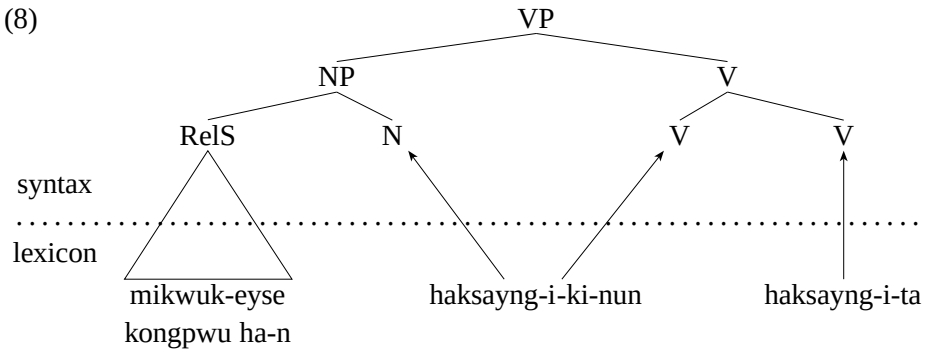
(7) a. ku salam-i papo-ka ani-ki-nun papo-ka ani-ta
that person-NOM fool-NOM NCOP-NMLZ-TOP fool-NOM NCOP-DECL
'It is true that that person is not a fool.' (VP-topic)

b. ku salam-i papo-ka ani-ki-nun ani-ta
that person-NOM fool-NOM NCOP-NMLZ-TOP NCOP-DECL
'That person is not a fool (but he is not so smart).' (ECC)

[2]The facts are subtle because prosodic prominence on the marker *-nun* can also trigger a negative implicature due to its contrastive properties, but we believe that speakers can have clear intuitions about ECC vs. VP-Topic constructions.

Although the details will come later, the structure we assign to (6)a is given in (8), where the two sister V nodes over to the right constitute the ECC. The upward arrows are explained in section 5.

(8)



As is clear from (8), the combination of noun+copula is a lexical unit – it is a single item in the lexicon.

3. Noun Copying and Verb Copying

Now we introduce the fourth construction mentioned in (4). Jo (2004a,2004b) discusses pairs of examples apparently involving the ECC and the copula, based on the simple example in (9)a:

(9) a. chelswu-ka pwuca-i-ess-e
chelswu-NOM rich-COP-PAST-DECL
'Chelswu was rich (a rich man).'

b. chelswu-ka pwuca-nun pwuca-i-ess-e
chelswu-NOM rich-TOP rich-COP-PAST-DECL

c. chelswu-ka pwuca-i-ki-nun pwuca-i-ess-e
chelswu-NOM rich-COP-NMLZ-TOP rich-COP-PAST-DECL

Jo argues that the relation between (9)b and (9)c shows that what is copied is either the N before the copula, for (9)b, or a larger constituent consisting of N and the copula, for (9)c, both coming from the same source in a transformational derivation involving the 'copy theory' of movement.[3]

There are several pieces of evidence which show that although (9)c is indeed an instance of the ECC, (9)b is not. Rather, (9)b is an 'N-Copy Construction' (which we will call 'NCC'), which reinforces the meaning of the N, and we translate it

[3]Strictly, Jo argues that a sequence of the subject plus some part of the predicate is copied, with the subject deleted in the second copy (Jo, 2004b, 172ff.).

roughly as 'truly'.[4] In other words, (9)b involves N copying, while (9)c involves V copying, as the underlining above suggests.

As mentioned above, the pragmatic hallmark of the ECC is that it sets up a negative implicature, without any assistance from other morphemes in the clause which may have adversitive or concessive meanings. This distinguishes (9)b from (9)c, and identifies only c as the ECC. While they both involve copying constructions (which will be related, but not identical, in our analysis), the key difference is that (9)b involves copying Ns, while (9)c involves copying Vs, and only the latter type has the negative implicature.

One clear difference between the two constructions can be observed from the alternation with the HCC. With noun and copula, the ECC alternates with the HCC (see (5)b), while the NCC does not:

(10) NCC does not alternate with HCC; ECC does alternates with HCC:

a. chelswu-ka pwuca-nun pwuca-i-ess-e (NCC)
chelswu-NOM rich-TOP rich-COP-PAST-DECL
'Chelswu is a truly rich man.'

b. *chelswu-ka pwuca-nun hay-ss-e (*HCC)
chelswu-NOM rich-TOP do-PAST-DECL
'Chelswu is a truly rich man.'

c. chelswu-ka pwuca-i-ki-nun pwuca-i-ess-e (ECC)
chelswu-NOM rich-COP-NMLZ-TOP rich-COP-PAST-DECL
'Chelswu is a rich man, but . . . .'

d. chelswu-ka pwuca-i-ki-nun hay-ss-e (HCC)
chelswu-NOM rich-COP-NMLZ-TOP do-PAST-DECL
'Chelswu is a rich man, but . . . .'

The HCC is clearly a V-V complex predicate, so its failure to work with a purely nominal first part as in (10)b is expected.

Next, the interaction with the negative copula is once again telling. From the simple example in (11)a, we might expect the following alternatives to be acceptable, which involve (respectively) copying the verb, copying the noun and verb, or just copying the noun:

(11) ECC can be negated; NCC can not:

a. chelswu-ka pwuca-ka ani-ta
chelswu-NOM rich-NOM NCOP-DECL
'Chelswu is not a rich man.'

[4]The 'prototypical' nature of the interpretation is also discussed for a copying construction in English by Ghomeshi et al. (2004), a construction they term 'contrastive focus reduplication'. They propose that the semantics of reduplication in English involves contrastive focus along the dimension of "PROTOTYPICAL/EXTREME/SALIENT".

b. chelswu-ka pwuca-ka ani-ki-nun ani-ta (negative ECC)
chelswu-NOM rich-NOM NCOP-NMLZ-TOP NCOP-DECL
'Chelswu is not a rich man (but he is very generous).'

c. chelswu-ka pwuca-ka ani-ki-nun
chelswu-NOM rich-NOM NCOP-NMLZ-TOP
pwuca-ka ani-ta (negative VP-topic)
rich-NOM NCOP-DECL
'As for not being rich, Chelswu is not rich.'

d. *chelswu-ka pwuca-nun pwuca-ka ani-ta (negative NCC)
chelswu-NOM rich-TOP rich-NOM NCOP-DECL
(int.) 'Chelswu is truly (not) rich.'

However, the last example is unacceptable, showing that while the ECC sets up a negative implicature, the NCC involves N copying and reinforces the positive property of the N. This explains why (11)d is strange – the N-copy part sets up a strong positive assertion, but then the verb negates it. Each example in (11)b–d has a different kind of interpretation, which shows that they cannot come from a common source.

Further differences exist between the ECC and the NCC. Delimiters like *-man* ('only') can be used as the marker on the first verb in the ECC but not in the NCC, as seen in (12):[5]

(12) ECC allows delimiters on the first predicate other than *-nun*; NCC does not:

a. chelswu-ka pwuca-i-ki-nun/man pwuca-ya (ECC)
Chelswu-NOM rich-COP-NMLZ-TOP/only rich-DECL
'Chelswu is rich/only rich, but . . . '

b. chelswu-ka pwuca-nun/*man pwuca-ya (NCC)
Chelswu-NOM rich-TOP/only rich-DECL
'Chelswu is truly (*only) rich.'

Finally, a proper noun cannot occur in the NCC at all:

(13) The predicate in the ECC can be a proper name, but not in the NCC:

a. ku salam-i John-i-ki-nun John-i-ya (ECC)
that person-NOM John-COP-NMLZ-TOP John-COP-DECL
'That person is John, but . . . '

b. ?*ku salam-i John-un John-i-ya (NCC)
that person-NOM John-TOP John-COP-DECL
'That person is truly John.'

[5]The copula *-i-* in (12)a is phonologically elided due to the following high [y].

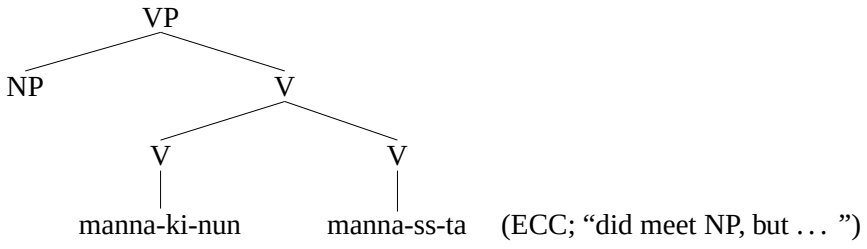
In summary, there are many differences between copying a noun (in the NCC) and copying a noun+copula unit (in the ECC), indicating two separate but related constructions. The noun+copula unit behaves like a verb, as predicted by the lexical analysis of the copula, but not by the clitic analysis.

4. Analysis of the Constructions

In this section we present our analysis of ECC, HCC and NCC constructions discussed above, with the exception of the Lexical Sharing analysis of the copula, which is given in the following section.

In the constructional approach, constructions are generalized types which express partial information that a surface configuration may inherit from. The structures that we need to account for are given in (14)–(16).

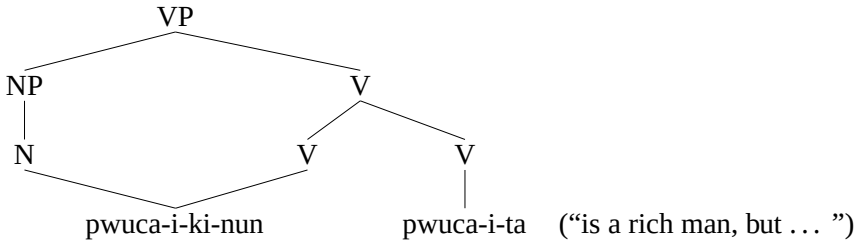
(14) ECC/HCC



The ECC involves copying a verb stem (*manna-* (‘meet’) in this example), while the HCC presents the content verb as complement to *ha-*. As far as syntactic structure is concerned, the constructions are identical.

Now if the verb in the ECC is a copularized noun, we have the structure in (15), presupposing the lexical sharing analysis to follow:

(15) ECC with copula

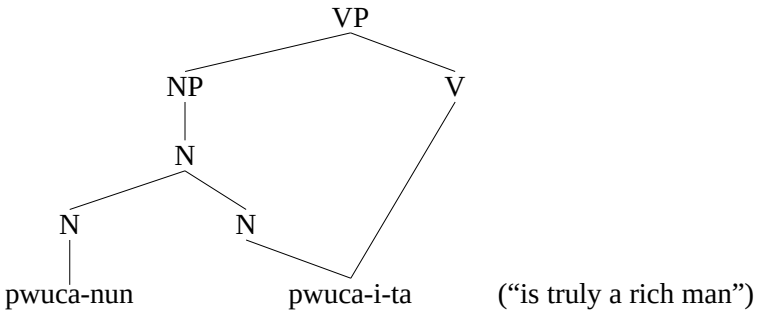


As this is the ECC, it is the V node which is intuitively copied; it just happens that the the first V participates in lexical sharing with a preceding N, which heads an NP. The regular principles of ordering place the NP in front of its sister V, and

by the Homomorphic Lexical Integrity property of the Lexical Sharing theory (see Wescoat (2002)), the N head of NP must (immediately) precede the left V copy. Informally, Lexical Sharing does not allow tangled trees, and so two nodes which share the same lexical item must be adjacent pre-terminals in the structure.

The structure in (15) contrasts with the NCC in (16), in which it is the N which is copied, and the second copy is lexically shared with a single V node:

(16) NCC with copula



Intuitively, the ECC and NCC involve copying the V or N stem, a stem which may be subject to further affixation (such that the non-stem parts of the two copies differ in form). We will adopt an approach here which copies on the basis of fine-grained semantic similarity, rather than surface phonological form, using Minimal Recursion Semantics (MRS). To begin, the semantics of the inflected verb *ilk-ess-ta* (‘read-PAST-DECL’) is as follows:

(17)
$$\begin{bmatrix} \textsc{phon} \left\langle \textit{ilk-ess-ta} \right\rangle \\ \textsc{syn} \begin{bmatrix} \textsc{head} \textit{ verb} \\ \textsc{arg-st} \left\langle \textsc{np}[\textsc{index}\ i], \textsc{np}[\textsc{index}\ j] \right\rangle \end{bmatrix} \\ \textsc{sem} \begin{bmatrix} \textsc{mode} \textit{ proposition} \\ \textsc{index} \boxed{4} s0 \\ \textsc{key} \boxed{5} \begin{bmatrix} \textsc{pred} \textit{ read} \\ \textsc{arg}0\ \boxed{4} \\ \textsc{arg}1\ i \\ \textsc{arg}2\ j \end{bmatrix} \\ \textsc{rels} \left\langle \boxed{5}, \begin{bmatrix} \textsc{reln} \textit{ temp-precede} \\ \textsc{arg}0\ \boxed{4} \\ \textsc{arg}1\ \textit{now} \end{bmatrix} \right\rangle \end{bmatrix} \end{bmatrix}$$

The semantics is stated as a list of *relations*, one of which is picked out as the KEY, and this will be the basis of the copying. The MODE feature represents the semantic mode of the expression such as *proposition*, *question*, *directive*, expressed by

the final MOOD marker in Korean. The preceding past tense marker *-ess-* provides the information in the second element in the RELS list. The information associated with the KEY comes directly from the lexical root, the V-root *ilk-* in (17).

We can now state the copy construction in (18), which forms copies based on the KEY in the MRS semantic form:

(18) *copy-cx*:

$$\left[\ \right] \rightarrow \begin{bmatrix} \text{SEM} \mid \text{KEY } \boxed{1} \\ \text{BAR } 0 \end{bmatrix}, \ \mathbf{H}\begin{bmatrix} \text{SEM} \mid \text{KEY } \boxed{1} \\ \text{BAR } 0 \end{bmatrix}$$

The specification of [BAR 0] allows for the presence of a restricted kind of adverbial in the copies; such adverbials are X^0s which are themselves modifiers of V^0 (see Sells (1998)). An identical adverbial must appear before each verb, so (19)b is unacceptable, contrasting with (19)a; and both verbs must match too, as (19)c shows.

(19) a. John-i Tom-ul cacwu manna-ki-nun cacwu manna-ss-ta]
John-NOM Tom-ACC often meet-NMLZ-TOP often meet-PAST-DECL
'John often met Tom, but . . . .'

b. *John-i Tom-ul cacwu manna-ki-nun congcong manna-ss-ta
John-NOM Tom-ACC often meet-NMLZ-TOP often meet-PAST-DECL

c. *John-i sinmwun-ul cacwu sa-ki-nun
John-NOM newspaper-ACC often buy-NMLZ-TOP
cacwu ilk-ess-ta
often read-PAST-DECL
'John often bought/read a newspaper, but . . . .'

Our proposal in terms of the KEY predicts these facts. In (19) the KEY will be the semantics of the adverbial 'often' which will directly take the verb's *relation* as its argument. Only in (19)a do the KEY values fully match as (18) specifies.

The ECC inherits from (18) and from a semantic constraint setting up a constrastive focus, and hence a negative implicature, given in (23) below. The NCC also inherits from (19) and from a semantic constraint expressing a reinforced positive assertion of the property denoted by the copied N (in (24) below).

With regard to the precise form of the copies, the first copy in the NCC is marked with *-nun*, while a verb in the ECC is first nominalized with *-ki* (which we treat via a FORM feature – see (23) below) before hosting *-nun*, or some other particle (the variation is shown in the examples in (12) above). We do not attempt to present a full account of the morphology in this paper (further details are given in Cho et al. (2004)).

The HCC is a complex predicate, a type of *hd-word-ph(rase)*, in which a head selects a complement:

(20) *hd-word-ph*:

$$\left[\text{BAR } 0\right] \rightarrow \boxed{1}\left[\text{BAR } 0\right], \mathbf{H}\begin{bmatrix}\text{BAR} & 0 \\ \text{COMPS} & \langle \ldots, \boxed{1}[(\text{FORM } \mu)], \ldots \rangle\end{bmatrix}$$

This general type of combination lets one X^0 select for another X^0 (including an optional FORM specification μ on the selectee), creating a new X^0. Complex predicates in Korean involve possibly recursive combinations of X^0 elements (see Sells (1998)).

Finally, any verbal complex predicate inherits from the type *hd-word-ph*, and every verb-headed phrase from *verb-headed-ph*:

(21) *verb-headed-ph*:

$$[\quad] \rightarrow \ldots \mathbf{H}\begin{bmatrix}\textit{verb} & \\ \text{BAR} & 0 \\ \text{TENSE} & \textit{value} \\ \text{MOOD} & \textit{value}\end{bmatrix}$$

This type requires that a lexical head should be specified for a value for TENSE and MOOD features, as appropriate for the type *verb*.[6]

Given that the HCC is a type of complex predicate, we need not assume any other constraints on the type of *hcc-ph* other than the lexical constraints on the auxiliary verb *ha-*, and the information inherited from *contrast-hd-ph*, which the ECC also inherits from. It is well-known that an initial *-(n)un*-marked phrase in Korean marks a Topic in the simplest sense, while sentence-internal *-(n)un* marks a contrastive phrase of some kind. Following Vallduví and Vilkuna (1998), we assume that any phrase may be specified as Topic and Focus and also may be independently specified as Constrastive (or not). As it is marked with *-(n)un* or some other suitable delimiter, and may also bear a phonological accent, it is the content verb in a copy construction which receives a Contrastive interpretation. Lee (2000) has argued that the negative implicature arises from a constructional meaning of 'Contrastive Topic'. We take the meaning of Contrastive Topic to be as summarized informally in (22):

(22) Contrastive Topic: The proposition is asserted, and that assertion implicates that there is at least one alternative proposition which is either false or whose truth value is not known (based on Oshima (2002)).

We then represent *contrast-hd-ph* as follows:

(23) *contrast-hd-ph*:

$$\left[\ \right] \rightarrow \begin{bmatrix}\text{PRED } \textit{contrast-topic} \\ \text{ARG}0\ \textit{s0}\end{bmatrix}, \quad \mathbf{H}\left[\ \right]$$

[6]Korean verbs are also specified as honorific (usually with the suffix *-(u)si-*) or non-honorific (unmarked); for simplicity, we omit consideration of honorification here.

This constraint specifies the construction *contrast-hd-ph* has a Contrastive Topic interpretation applied to the situation index of the predication of the head verb.

The contribution of 'reinforcement of prototype' (see footnote 4) for the NCC is likewise specified in (24):

(24) *reinforce-hd-ph*:

$$\left[\textsc{sem}\left[\textsc{rels}\left\langle \ldots, \left[\textsc{pred}\ \textit{prototypical}\right], \ldots \right\rangle\right]\right] \rightarrow$$

$$[\textsc{index}\ \boxed{1}\ \textit{indiv}]\ ,\ \mathbf{H}[\textsc{index}\ \boxed{1}\ \textit{indiv}]$$

In sum, the surface constructions that we have discussed here inherit as follows, which graphically illustrates the commonalities:

(25)

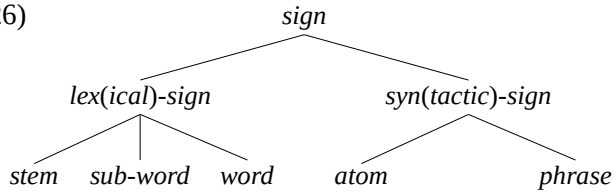
ecc-ph:	*copy-cx*	$\wedge$	*contrast-hd-ph*	$\wedge$	*vb-headed-ph*			
hcc-ph:			*contrast-hd-ph*	$\wedge$	*vb-headed-ph*	$\wedge$	*hd-word-ph*	
ncc-ph:	*copy-cx*	$\wedge$	*reinforce-hd-ph*					

5. Lexical Sharing Analysis

5.1. Lexical Sharing in HPSG

Wescoat (2002) argues that the atomic units of phrase structure are neither words, as claimed by Di Sciullo and Williams (1987), nor morphemes, as assumed in Autolexical Syntax (see Sadock (1991)), but rather lexical-category-bearing *atomic constituents*, each of which maps to a word, which *instantiate*s the atomic constituent. The basic idea of lexical sharing is then that a single word may instantiate multiple atomic constituents. This scheme provides a straightforward model of words that appear to straddle a phrase boundary. Lexical sharing may be implemented using the basic machinery of HPSG, in which there is a basic sort of *sign*. Two subtypes of *sign*, namely *phrase* and *word*, have been traditionally employed for representing phrase-structure constituents; thus, standard HPSG is among those theories that regard words as the atoms of phrase structure. In the lexical sharing approach we divorce the type *word* from this role, and have a new, properly syntactic type to represent atomic constituents in phrase-structure, namely *atom*. The modified *sign* hierarchy is shown in (26).

(26)



The type of an AVM determines what features and what types of values the AVM may contain. The principal new type declarations are given in (27).

(27) a.
$$\begin{bmatrix} \textit{lex-sign} \\ \text{PHON(OLOGY)} \quad \textit{phon} \\ \text{INST(ANTIATE)S} \quad \textit{nelist}(\textit{synsem}) \end{bmatrix}$$

b.
$$\begin{bmatrix} \textit{syn-sign} \\ \text{YIELD} \quad \textit{nelist}(\textit{phon}) \\ \text{SYNSEM} \quad \textit{synsem} \end{bmatrix}$$

c.
$$\begin{bmatrix} \textit{atom} \\ \text{ARG-ST} \quad \textit{list}(\textit{synsem}) \\ \text{YIELD} \quad \langle [\ ] \rangle \end{bmatrix}$$

d.
$$\begin{bmatrix} \textit{phrase} \\ \text{D(AUGH)T(E)RS} \quad \textit{nelist}(\textit{syn-sign}) \end{bmatrix}$$

The basic intuition is that the *word* is the exponent of the syntactic *atom*. Hence, in the declarations above, on the one hand, a lexical sign has a PHON value, like a word, and on the other hand it instantiates the SYNSEM of an *atom*, which is a syntactic sign. An atom is one type of syntactic sign, and every syntactic sign has a YIELD feature, whose value is a list of PHON values. The difference between PHON and YIELD is explained more below. For an atom, the YIELD list is of length one; for a phrase, the list of PHON values in YIELD will represent the order of constituents.[7]

The feature INSTS in (27)a implements lexical sharing: every *word* contains, as the value of INSTS, an ordered list enumerating each *atom* that the *word* instantiates (see (28)). The PHON value of the word becomes a member of the list value of the atom's YIELD by the following constraint. In the normal case, when there is no actual lexical 'sharing', the INSTS list is simply of length one.

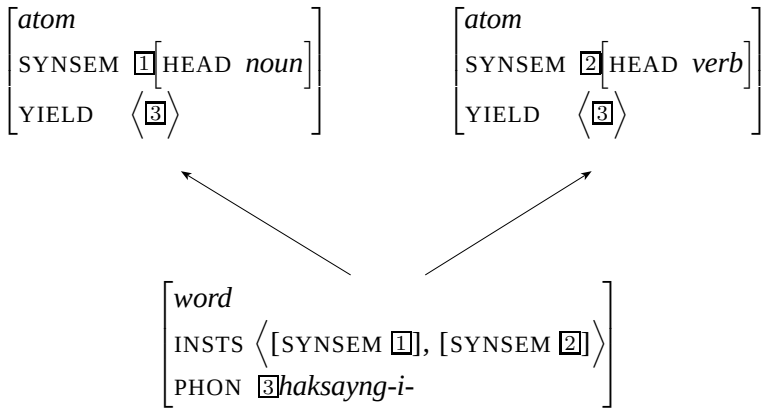
(28)
$$\textit{word} \Rightarrow \begin{bmatrix} \text{INSTS} \left\langle \begin{bmatrix} \textit{atom} \\ \text{SYNSEM} \ \boxed{1} \\ \text{YIELD} \quad \langle \boxed{3} \rangle \end{bmatrix}, \ldots, \left(\begin{bmatrix} \textit{atom} \\ \text{SYNSEM} \ \boxed{2} \\ \text{YIELD} \quad \langle \boxed{3} \rangle \end{bmatrix} \right) \right\rangle \\ \text{PHON} \ \boxed{3} \end{bmatrix}$$

[7] The feature YIELD is different from the feature DOMAIN in several respects. For example, while the mother's DOMAIN value is the concatenation of the daugther's DOMAIN values, the mother's YIELD value is the realization of the PHON value(s). See Kathol (2000) for detailed discussion of the DOMAIN feature.

This says that each word has a PHON value and it instantiates some number of atoms, each of which has the word's PHON value as its YIELD.

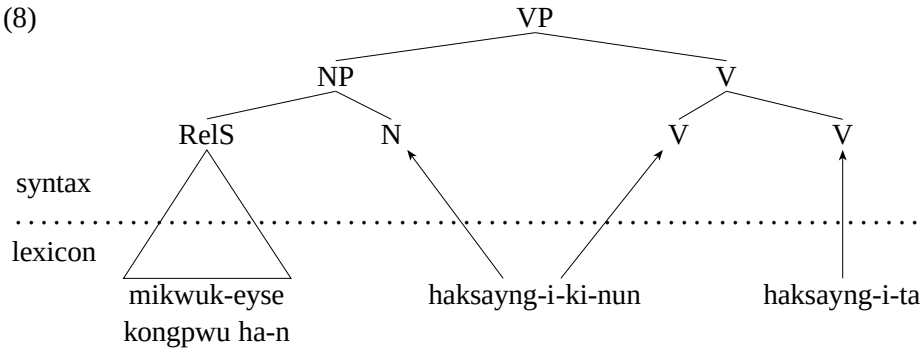
The effect of (28) is illustrated by the schematization in (29) of an instance of lexical sharing (the form *haksayng-i-*):

(29)



Let us compare (29) with (8), repeated here; we can see that (29) correctly captures the lexical sharing for the word *haksayng-i-ki-nun*. The last word *haksayng-i-ta* just instantiates one atom, as we describe below.

(8)

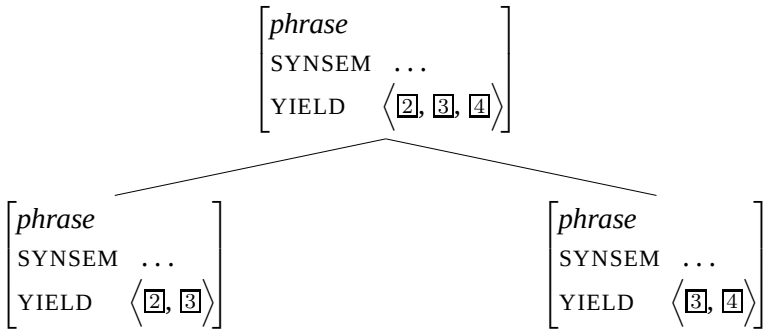


To complete the implementation of lexical sharing, the new list-valued feature YIELD is strictly speaking not just the concatenation of the elements on a list, but it is the *uniq* of a list. We state *uniq* as a constraint on the type *phrase*:

(30) $$\textit{phrase} \Rightarrow \begin{bmatrix} \textsc{yield} & \textit{uniq}(\boxed{1} \oplus, \ldots, \oplus \boxed{n}) \\ \textsc{dtrs} & <[\textsc{yield}\,\boxed{1}], \ldots, [\textsc{yield}\,\boxed{n}]> \end{bmatrix}$$

uniq is a function on lists, which contracts a list to a list of unique members, by removing the second of two adjacent identical occurrences of member μ. If there are two identical occurrences of member μ which are non-adjacent, *uniq* is undefined. In the case of lexical sharing, a word whose PHON value is $\boxed{3}$ (cf. (29)) will participate in a structure where *uniq* applies to lists of $\langle\, \boxed{2}, \boxed{3} \,\rangle$ and $\langle\, \boxed{3}, \boxed{4} \,\rangle$:

(31)

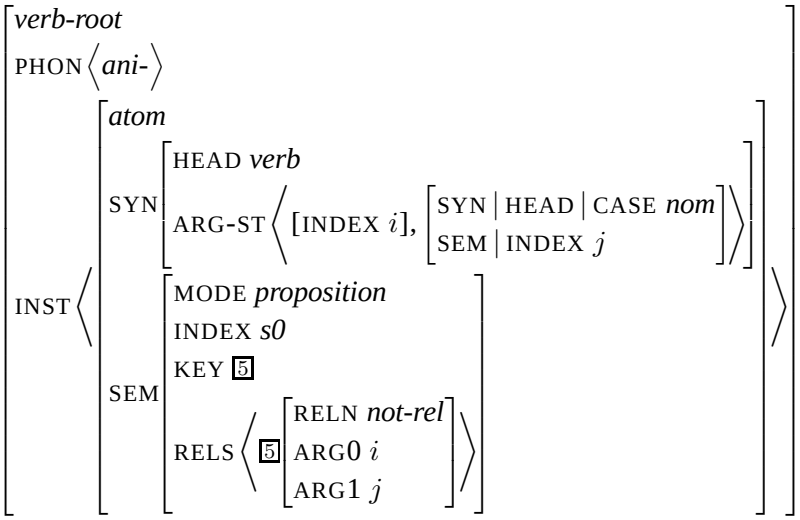


The yield of the phrase is ⟨ 2, 3, 4 ⟩ and not ⟨ 2, 3, 3, 4 ⟩.

5.2. The Copula

Finally, we come to the analysis of the copula itself, which will be easier to see if we start with the negative copula first. The entry for this form is shown in (32) (we assume a sort *verb-root* for the basic representation of verb roots in Korean):

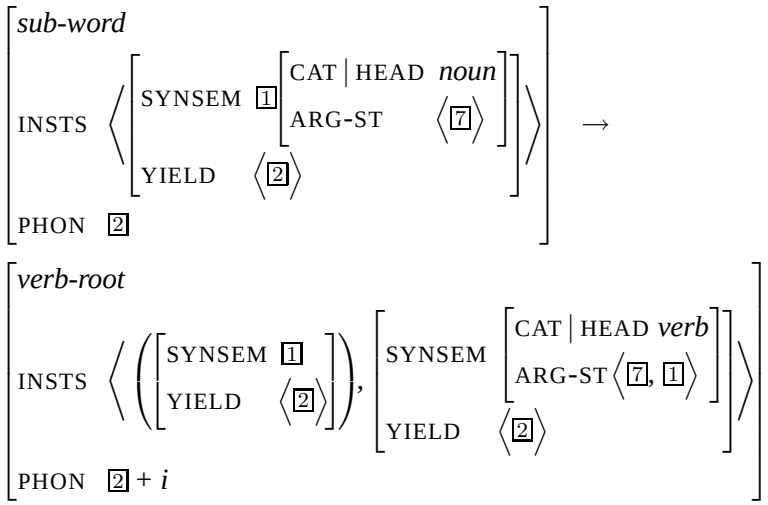
(32)



This root will be the basis of a word which instantiate a verb atom, which itself selects for two NPs, the second of which is specified to be in the nominative case.

Unlike the negative copula, the positive copula *-i-* does not exist as a root itself; it is intuitively an affix. We provide a lexical rule which takes a noun *sub-word* as input, and returns a *verb-root*. This new lexical form instantiates two syntactic atoms, as shown in (32):

(33) Copularization:



The output form instantiates two atoms in the syntax, an N (which heads NP) and a V (which heads VP), and may be input to further lexical rules and inflection. Hence, this is appropriate for the form *haksayng-i-ki-nun* in (8). The lexical rule puts the relevant syntax and semantics of the host N as information about the second argument of the V that the output form instantiates. Nevertheless, this is still a two-place V, an atom which will eventually combine in syntax with a complement NP and with a subject NP (3). The parenthesis in the rule around the first element on the INSTS list allows the word created by positive copularization to optionally instantiate one just one atom in syntax, a V, as for *haksayng-i-ta* at the end of (8).

6. Conclusion

The facts from the ECC show that noun plus copula is a lexical unit in Korean; they also show that the copula itself has no independent status as a verb. We presented an overview of the ECC and some related constructions, in particular the noun-copying NCC, which has different semantics from the ECC. Finally, to deal with the fact that noun plus copula is one lexical unit, but corresponds to both an N node and a V node in syntax, we adopted the Lexical Sharing approach of Wescoat (2002) within HPSG.

References

Aoyagi, Hiroshi. 2005. On Predicate Focus Constructions in Korean and Japanese. In S. Kuno et al. (ed.), *Harvard Studies in Korean Linguistics*, volume 11, Harvard University: Dept. of Linguistics, to appear.

Cho, Sae-Youn, Kim, Jong-Bok and Sells, Peter. 2004. Contrastive Verb Constructions in Korean. In S. Kuno et al. (ed.), *Harvard Studies in Korean Linguistics*, volume 10, pages 360–371, Harvard University: Dept. of Linguistics.

Cho, Young-mee Yu and Sells, Peter. 1995. A Lexical Account of Inflectional Suffixes in Korean. *Journal of East Asian Linguistics* 4, 119–174.

Choi, Kiyong. 2003. The Echoed Verb Construction in Korean: Evidence for V-raising. In Patricia M. Clancy (ed.), *Japanese/Korean Linguistics*, volume 11, pages 457–470, CSLI, Stanford Linguistics Association.

Di Sciullo, Anna-Maria and Williams, Edwin. 1987. *On the Definition of Word*. Cambridge: MIT Press.

Ghomeshi, Jila, Jackendoff, Ray, Rosen, Nicole and Russell, Kevin. 2004. Contrastive Focus Reduplication in English (The Salad-Salad Paper). *Natural Language and Linguistic Theory* 22, 307–357.

Jo, Jung-Min. 2004a. Variation in Predicate Cleft Constructions in Korean: Epiphenomena at the Syntax-PF Interface. In S. Kuno et al. (ed.), *Harvard Studies in Korean Linguistics*, volume 10, pages 424–437, Harvard University: Dept. of Linguistics.

Jo, Jung-Min. 2004b. *Grammatical Effects of Topic and Focus Information*. Ph. D.thesis, University of Illinois at Urbana-Champaign.

Kathol, Andreas. 2000. *Linear Syntax*. New York: Oxford University Press.

Kim, Jong-Bok and Chung, Chan. 2002. Korean Copula Constructions: A Construction and Linearization Perspective. *Ene* 27, 171–193.

Kim, Jong-Bok, Sells, Peter and Wescoat, Michael T. 2004. Korean Copular Constructions: A Lexical Sharing Approach. In M. E. Hudson, S.-A. Jun and P. Sells (eds.), *Japanese/Korean Linguistics*, volume 13, CSLI, Stanford Linguistics Association, to appear.

Kim, Youngsun. 2002. The Syntax of Korean Verbal Focus Constructions. In *Proceedings of the 2002 LSK International Summer Conference*, pages 333–344, The Linguistic Society of Korea.

Koopman, Hilda. 1984. *The Syntax of Verbs*. Dordrecht: Foris Publications.

Lee, Chungmin. 2000. Contrastive Predicates and Conventional Scales. In A. Okrent and J. Bolye (eds.), *CLS*, volume 36, pages 243–257, Chicago.

Nishiyama, Kunio and Cho, Eun. 1998. Predicate Cleft Constructions in Japanese and Korean: The Role of Dummy Verbs in TP/VP-preposing. In Noriko Akatsuka et al. (ed.), *Japanese/Korean Linguistics*, volume 7, pages 463–479, CSLI, Stanford Linguistics Association.

Oh, Mira. 1991. The Korean Copula and Palatalization. *Language Research* 27, 701–724.

Oshima, David Y. 2002. Contrastive Topic As A Paradigmatic Operator. Ms. Stanford University (available at `http://www.stanford.edu/~davidyo`).

Sadock, Jerrold M. 1991. *Autolexical Syntax: A Theory of Parallel Grammatical Representations*. Chicago: University of Chicago Press.

Sells, Peter. 1997. Positional Constraints and Faithfulness in Morphology. In S. Kuno et al. (ed.), *Harvard Studies in Korean Linguistics*, volume 7, pages 488–503, Dept. of Linguistics, Harvard University.

Sells, Peter. 1998. Structural Relationships Within Complex Predicates. In Byung-Soo Park and James Yoon (eds.), *Proceedings of ICKL 11*, pages 115–147, Seoul: International Circle of Korean Linguists.

Vallduví, Enric and Vilkuna, Maria. 1998. On rheme and kontrast. In Peter Culicover and Louise McNally (eds.), *Syntax and Semantics 29: The Limits of Syntax*, pages 79–108, New York: Academic Press.

Wescoat, Michael T. 2002. *On Lexical Sharing*. Ph. D.thesis, Stanford University.

Yoon, James Hye-Suk. 2003. What the Korean Copula Reveals about the Interaction of Morphology and Syntax. In Patricia M. Clancy (ed.), *Japanese/Korean Linguistics*, volume 11, pages 34–49, CSLI, Stanford Linguistics Association.