

Korean Copular Constructions: A Lexical Sharing Approach

Jong-Bok Kim, Peter Sells and Michael T. Wescoat

Kyung Hee University, Stanford University, and University of California, Santa Cruz

1. Introduction

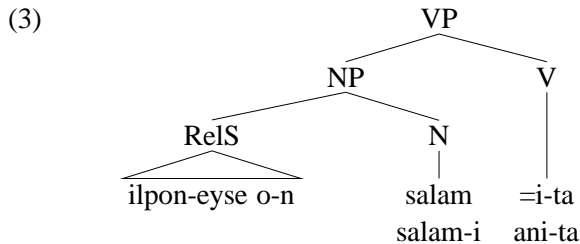
While the Korean copula *-i-* forms a phonological word with its preceding N host (see Oh (1991)), it also shows evidence of being an independent syntactic head which takes a phrasal complement, as in (1). The negative copula *ani-* shows a similar structure, but without the phonological cohesion.

- (1) a. ku haysayng-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 haysayng-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.’

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

In (1) the modifier *ilpon-eyse o-n* forms a constituent with the head noun *salam*. This N is also the host of the positive copula, and so a favored analysis has been that *-i-* is a clitic (e.g., Yoon (2003)) – a V in syntax that forms a phonological word with its adjacent N host, indicated by the ‘=’ in (3). We refer to this as the ‘clitic analysis’. The negative copula has the same syntax but is not a clitic.



We argue in this paper that while both copula forms do have a syntactic structure which allows for modification of the N *salam* inside NP, the clitic analysis of the copula is not correct. Rather, the copula forms a true lexical word with its N host, shown by three kinds of evidence:

- (4) a. the systematic absence of final suffixes on the complement N (as in (2)a);

- b. the phonological rules of Palatalization and Coda Neutralization which show that N+Copula is a lexical unit (see Cho and Sells (1995));
- c. a novel syntactic argument, coming from the Echo Contrastive Construction (ECC).

The copula-as-clitic analysis makes incorrect predictions for (4)b and (4)c, as it gives the wrong account of what the lexical units are, and leaves the correlation of lexical status with the restriction in (4)a as accidental.

The apparent conflict between the 3 lexical properties above and the structure in (3) with respect to *-i-* is in fact predicted by the Lexical Sharing approach of Wescoat (2002), a theory of the relation between words and the syntactic structures they instantiate, which is independently motivated for a variety of phenomena including English possessives (*'s*), contracted auxiliaries (e.g., *I'll*), and pronominal determiners (e.g., *[These] are good*), Hindi noun incorporation, Romance and Germanic preposition-determiner compounds, and second-word clitics.

This paper is organized as follows: in section 2 we briefly discuss different uses of the copula, and then sections 3-5 address each of the properties in (4) in turn. Section 6 presents the Lexical Sharing analysis of the copula that reconciles the syntactic structure in (3) with the properties in (4).

2. Types of the Copula

Yoon (2003) identifies two different uses of the Korean copula, the canonical predicational use, and the ‘inverse’, which intuitively puts the predicate initial and has the subject as the complement of the copula:

- (5) a. ku haysayng-un cinccalo [ilpon-eyse o-n salam]-i-ta (canonical)
that student-TOP actually [Japan-from come-MOD person]-COP-DECL
‘That student is actually a person from Japan.’
- b. ilpon-eyse o-n salam-un cinccalo [ku haksayng]-i-ta (inverse)
Japan-from come-MOD person-TOP actually [that student]-COP-DECL
‘The person from Japan is actually that student.’

In all, we believe that there are 4 types of copula construction in Korean, as in (6):

- (6) a. predicational copula construction (= canonical)
- b. specificational copula construction (= inverse)
- c. cleft construction
- d. fragment construction

The first 3 types have similar, though not identical, syntactic structures, while fragments are rather different.

2.1. Fragments

The fragment uses of the copula provide some apparent counterexamples to the claim (see (2) above) that final particles and the copula compete for the same position, and hence cannot cooccur. The following examples from Han (1996) and Sells (1997) illustrate:

- (7) a. yeki-eyse-to-i-nya b. ikes-to-i-nka
here-at-also-COP.PRES-Q this-also-COP.PRES-Q
‘Is it here also?’ ‘Is this it also?’

- (8) pwusan-eyse-nun-i-lkka
 Pusan-in-FOC-COP-Q
 ‘Even in Pusan?’

These fragments work best in interrogative environments (‘Is it X?’) and no example corresponds directly to a full clause with a missing argument. Rather, they are fragments which do not have a full sentential structure and which can only be used in specific contexts, in contrast to other uses of the copula which do not require any special context. In fact, the copula seems to be used in a way that depends pragmatically on the verb in a previous linguistic context, though the best literal translation of an example like (8) may just be ‘Is that the case even in Pusan?’, where *that* is anaphoric to something in the context. This context-dependence is illustrated by the following exchanges:

- (9) A: ecey manh-un kaswu-ka mayak-ulo caphi-ess-tay
 yesterday many-MOD singer-NOM drug-for arrested.PASS-PAST-be.reported
 ‘It is reported that many singers were arrested yesterday for drug use.’

- B: kulay, wuli-ka cohaha-nun ku kaswu-to(-i)-tela
 right we-NOM like-MOD the singer-also-COP-recollection
 ‘Right, the singer we like was also ...’

- (10) A: Mary-nun etten yuklyu-to mek-ci anh-nun-ta
 Mary-TOP any meat-kind eat-COMP NEG-PRES-DECL
 ‘Mary doesn’t eat any meat.’

- B: kyeylan-kkaci-to(-i)-ta
 egg-even-also-COP.PRES-DECL
 ‘(Not) even eggs.’

In the typology of lexical rules that we develop below, the fragment use of the copula is a word-to-word rule (see (28)c).

2.2. Full-Fledged Uses

Following the argumentation in Mikkelsen (2002) (see Heycock and Kroch (2002)), we will treat the predicational and specificational uses of the copula (canonical and inverse in Yoon (2003)) as having sufficiently different properties that one should not be derived directly from the other. In the examples below, we assume we have a notional subject *Sparky* and a notional predicate *haksayng* (‘student’); the predicational copula has the properties in (11). By ‘syntactic subject’ and ‘syntactic predicate’ we mean ‘first argument of *be*’ and ‘second argument of *be*’ (the complement N).

- (11) Predicational copula (no restrictions on Information structure):

Syntactic properties: *Sparky* = subject, *haksayng* = predicate

- | | |
|---|---|
| a. What is Sparky? (Sparky = Topic) | c. Who is the student? (student = Topic) |
| b. [Sparky] _T -nun [haksayng] _F -i-ta
‘Sparky is [a student] _F .’ | d. [Sparky] _F -ka [haksayng] _T -i-ta
‘[Sparky] _F is the student.’ |

The fact that both information structure profiles ([T-F] and [F-T]) are possible with this syntactic configuration shows that the canonical copula structure places no information structure restrictions on its argument positions, and that the linking to syntax simply follows the argument-structure hierarchy.¹

The specificational copula is rather different, and only allows the [T-F] order on the syntactic constituents (which are intuitively ‘reversed’):

- (12) Specificational copula (predicate is Topic, subject is Focus):
- Syntactic properties: *Sparky* = predicate, *haksayng* = subject
- | | |
|--|--|
| <p>a. What is Sparky? (Sparky = Topic)</p> <p>b. #[haksayng]_T-un [Sparky]_F-i-ta
 ‘The student is [Sparky]_F.’</p> | <p>c. Who is the student? (student = Topic)</p> <p>d. [haksayng]_T-un [Sparky]_F-i-ta
 ‘The student is [Sparky]_F.’</p> |
|--|--|

As noted by Yoon (2003), there is no true inverted predication in Korean (the same as English):²

- (13) #haksayng-i Sparky-i-ta

The unacceptability of this example shows that Korean has no simple inverse structure where the intuitive predicate *haksayng* appears as the surface subject and *Sparky* is the surface predicate. Rather, the construction must be used for the specific information-structure purposes given in (13).

The cleft use of the copula is like the specificational use in imposing obligatory [T-F] information structure (see e.g., (20)a below), but otherwise it has the syntactic properties of the canonical copula (for example, it allows predicate doubling as in (23)b below).

3. Lexical Status I: Restrictions on Final Particles

In this section we concentrate on the contrast between the examples in (2):

- (2) a. apeci-nun hakca(*-ka/*-nun/*-to)-i-ta
 father-TOP scholar(*-NOM/*-FOC/*-FOC)-COP-DECL
- b. apeci-nun hakca(-ka/-nun/-to) ani-ta
 father-TOP scholar(-NOM/-FOC/-FOC) NCOP-DECL

As far as we can determine, the restriction illustrated by (2)a holds for all ‘full-fledged’ uses of the copula. To account for this restriction, Yoon (2003) proposes that the positive copula takes a complement which is marked for Null Case (which has \emptyset realization), but which is a case alongside Nominative or Accusative. An NP with Null Case is not the same (at all) as an NP which lacks a case marker at all (where case has been ‘dropped’). In a similar analysis, Kim and Chung (2002) stipulate that the complement of the copula cannot be specified for structural case, giving the same effect as Yoon’s Null Case, by blocking other structural cases.

However, it is an accident on these approaches that it is precisely the form of the copula which cannot stand alone as a syntactic word which is also the form that takes a Null Case complement. The lexicalist analysis attempts to link these two facts. In our view, the negative copula *ani-ta* takes a phrasal complement and as (noun) phrases have case in Korean, this complement has case. That case is the default Nominative, as the special semantic conditions which license Accusative or Dative are not met. In contrast, the lexical attachment of the positive copula *-i-ta* prevents case from appearing, without appeals to unmotivated values for a case feature.

¹As in English, the [F-T] information structure requires accent on the focal element.

²Though this fact leads Yoon and us to draw different conclusions about the relation between the predicational and specificational uses of the copula.

3.1. Korean Nominal Morphology and the Copula

Based on the classification of Yang (1972), Cho and Sells (1995) adopt the well-known template in (14) for the Korean nominal system, with each position exemplified by the particles shown in (15) and (16).³

(14) N-stem – Postposition – Conjunctive – Delim (X-LIM) – Delim (Z-LIM)

(15)	Postpositions		Conjunctives
	<i>eykey(se), hanthey(se)</i>	dative	<i>hako, (k)wa</i> conjunctor
	<i>ey, eyse</i>	locative	<i>pota</i> comparator
	<i>ey, (u)lo</i>	directive	<i>(i)na</i> disjunctor/ 'something like'
	<i>(u)lo</i>	instrumental	
	<i>kkaci</i>	goal	<i>pwuthe</i> 'from'
	<i>hako, (k)wa</i>	comitative	<i>chelem</i> 'like'
	<i>kkey</i>	dative (hon.)	
	<i>kkeyse</i>	hon. subj.	

(16)	Delimiters	
	'X-LIM'	'Z-LIM'
	<i>man</i> 'only'	<i>(n)un</i> TOPIC/FOCUS
	<i>kkaci</i> 'even'	<i>i/ka</i> NOM
	<i>mace</i> 'even'	<i>(l)ul</i> ACC
	<i>cocha</i> 'even'	<i>uy</i> GEN
	<i>pakkey</i> 'only'	<i>to</i> 'also'/FOCUS

The canonical predicative copula does not allow any Z-LIM element on the host N, while all other suffixes are possible, in principle.⁴

Nominative and Accusative case markers appear on NP arguments, but in Korean these case markers may be 'supplanted' by other particles indicating emphasis or focus. The phrasal complement to the negative copula *ani-* is no different in this regard. In (17), the complement to the negative copula bears the particle *-to*, meaning 'also'.

(17) ku i-nun uysa-to ani-ko umak-ka-to ani-ta
 that one-TOP doctor-also NEG.COP-CONJ music-person-also NEG.COP.PRES-DECL
 'He is neither a doctor nor a musician.'

However, it is not possible to attach *-to* directly to the predicate nominal before the positive copula, as in (18):

(18) *... umak-ka-to-i-ta
 ... music-person-also-COP.PRES-DECL
 '... is also a musician.'

Instead, if a Z-LIM like *-to* is to cooccur with the positive copula, it is necessary use the periphrastic construction in (19).

³We present the template here for illustrative purposes, while fully recognizing that its descriptive effects should follow from more principled morphological constraints (see Sells (1997), Yoon (2002)), and also section 3.3.

⁴The same facts are true for the Japanese copula *-da*, and the corresponding 'Z-LIM' elements in Japanese are: *-wa* ('TOPIC/FOCUS'), *-ga* (NOM), *o* (ACC), *no* (GEN).

- (19) ku i-nun uysa-i-ki-to ha-ko umak-ka-i-ki-to ha-ta
 that one-TOP doctor-COP-NOMIN-also do-CONJ music-person-COP-NOMIN-also do.PRES-DECL
 ‘He is both a doctor and a musician.’

Note that the explanation for (18) cannot involve some problem with *-to* taking scope within the complement of the copula, as the same semantic structure obtains in (17), which is acceptable.

The same contrast (anything but a Z-LIM) shows up in (20), in a cleft use:

- (20) a. ilen iyaki-ka cekyong toy-nun kes-un swuni-hanthey-man-i-ta
 this.kind story-NOM apply PASS-PRES.MOD NOMIN-TOP Sooni-DAT-only-COP.PRES-DECL
 ‘It is only to Sooni that this story is applicable.’
- b. *... swuni-hanthey-man-un-i-ta
 ... Sooni-DAT-only-FOC-COP.PRES-DECL
 ‘It is only to Sooni ...’

Other examples of the contrastive difference between the positive and negative copulas can be found in Sells (1997) and Yoon (2003). The clear generalization then about the positive copula is that it effectively attaches in the last – Z-LIM – slot of the templatic structure in (14), and therefore any suffix that falls in that slot cannot cooccur with the copula. On the other hand, the negative copula takes a full syntactic complement, and so it shows no restrictions on suffixes.

3.2. Postpositions

Yoon (2003) claims that postpositions (see (15) may not be part of the N complement to the copula; this is predicted by the fact that the complement is assigned Null case, blocking any other case-marking suffix.⁵ However, this seems to be too strong a factual claim (example (21) from Kim and Chung (2002)):

- (21) kuke-n [poncilcekulo talun iyu-eyse]-i-ta.
 that-Top essentially different reason-at COP-DECL
 ‘That is essentially due to a different reason.’

Other examples illustrate the same possibility:

- (22) a. cwulpalcem-un Seoul-eyse-ya (< i-ta)
 starting point-top Seoul-from-COP
 ‘The starting point is (from) Seoul.’
- b. sihem pemwi-nun yeki-pwuthe-i-ta
 exam range-TOP here-from-COP-DECL
 ‘The exam range is from here.’

(23) presents some corpus examples, with a cleft use of the copula:

- (23) a. TV syophing-i sicak-toy-n kes-un 80-nyen-tay cwungpan-pwuthe-i-ta
 TV shopping-NOM start-become-MOD thing-TOP 80’s middle-from-COP-DECL
 ‘It is the middle 80’s when TV homeshopping began.’

⁵Actually, if Null case is a structural case, as it would seem to be, it ought to be able to ‘stack’ with semantic cases introduced by postpositions, (correctly) allowing examples such as those here as grammatical.

- b. TVsyophing-i sicaktoy-n kes-un, kuttay-pwuthe-nun
 TVshopping-NOM start-become-MOD thing-TOP that.time-from-TOP
 kuttay-pwuthe-i-ta
 that.time-from-COP-DECL
 ‘It is true that the time TV homeshopping began was then.’

Interestingly, the predicate part in (23)b is doubled, emphasizing the truth of the statement. Such doubling is only possible in the predicational and cleft uses of the copula, but not in the specificational use (observed by Yoon (2003)).⁶

3.3. Lexical Analysis

In summary, except for its fragment use, the copula takes an N complement, and that N may itself host any of the suffixes in (15)–(16) except for the Z-LIMs.

Sells (1997) and Yoon (2002) agree that the most important generalizations about nominal suffix ordering are those in (24):

- (24) a. The Postposition particles attach right after the N stem.
 b. The ZLIM particles are final within the word.

In a hierarchical lexicon that is the basis of any HPSG analysis, we set up different types corresponding to meaningful classes (sub-)lexical items. For nominals, we would start with the sorts in (25)a, which participate in suffixal processes as described in b–d:

- (25) a. Korean nominal sorts: *stem*, *sub-word*, *word* (see (40) below).
 b. Any sort can inherit directly as the next sort ‘up’, so no suffix is obligatory (for a noun) (see Yoon (1995)).
 c. Postposition maps a *stem* to a *sub-word*.
 d. ZLIM maps a *sub-word* to a *word*.

These will be expressed as exponent-creating lexical rules (see section 6 below), rather like the realizational approach to morphology of Stump (2001), which also has rules applying sequentially to successively add phonological exponents of morpho-syntactic information. Essentially, these rules just realize inflection on a noun, and the sorts give a familiar structure to a Korean word:

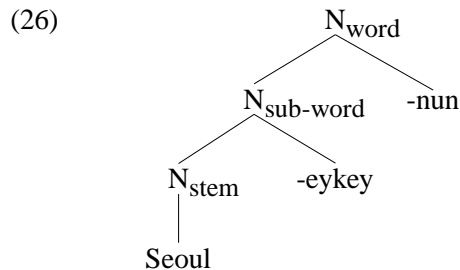
⁶Yoon (2003) observes that a postposition cannot precede the copula in examples like (i)a, while (i)b is fully acceptable.

- (i) a. *ku chayk-un selap-ey-i-ta
 b. ku chayk-un selap-ey iss-ta

However, as Korean has the specialized verb *iss-ta* for expressing existence and location, this form blocks *-i-ta*, which is the reason why (i)a is ungrammatical. We note that (i)c is also ungrammatical:

- (i) c. *ku chayk-un selap-ey ani-ess-ta
 d. ku chayk-un sela-ey eps-ess-ta

Hence, the constraint is not against having postpositions per se in the complement of the copula, but rather that Korean has specialized verbs for the ‘be’ of existence and location.



As noted above, any sort can directly inherit as the next sort ‘up’, so that just *Seoul* can also be classified as a *word* (and a *sub-word*) without any suffixation.

Korean grammar also involves some verbalizing rules which create forms like *N-kath-ta* (‘look like’) and *N-tap-ta* (‘be N-ish’), which are effectively derivational processes which take an *N stem* into the verbal part of the lexicon hierarchy (see (28) below). These verbalizing suffixes do not allow any of the nominal suffixes in (15)–(16) on the *N*, hence the restriction to a *stem*, in clear contrast to the copula:

- (27) a. *salam-man/kkaci kath-ta
 ‘looks only/even like a person’
- b. salam-man/kkaci-i-ta
 ‘is only/even a person’

We note that the differential sub-lexical selection of *-tap-ta* and *-i-ta* seems impossible to account on an account that takes them both to be syntactic heads which take phrasal complements. The relevant ‘derivational’ rules (in (28)) will create new verbal roots:

- (28) a. For *-kath-ta, tap-ta*: map N_{stem} to V_{root} .
- b. For copula *-i*: map $N_{\text{sub-word}}$ to V_{root} .

There are only two classes of lexical rules which change the type of the input element: those which take a *stem* as input, and those which take a *sub-word* as input. There are also lexical rules which preserve type. For instance, a semantic delimiter like *-man* can map a *stem* to a *stem* or a *sub-word* to a *sub-word*, though it never applies to a *word*, appearing outside a Z-LIM. Hence words like *Seoul-man-eykey* and *Seoul-eykey-man* are both predicted to be well-formed.

Korean does have some *word-word* rules, which are completely category-blind, like the spurious/copy marker *-tul* and the discourse-politeness marker *-yo*. These suffixes do not provide or add any information about the internal properties of the word they inflect, but rather about the larger context of use. As noted above, the fragment copula falls in this class, in the sense that it takes a word as input:

- (28) c. For fragment copula *-i*: map N_{word} to V_{root} .

So, in summary, the grammar gives the same explanation for the complementary distribution of the Nominative case marker *-i/ka* and the Topic/Focus marker *-(n)un* – both are the result of morphological exponent rules which apply to a *sub-word* to create a different object in the lexicon – and the complementary distribution of the copula *-i* and the Topic/Focus marker *-(n)un*.

4. Lexical Status II: Lexical Phonological Rules

Now we offer two further kinds of argument for the lexical status of N+copula. The first kind comes from phonology, showing that N+copula constitutes a lexical word, rather than two words making up a phonological word, which is the (phonological) result of the clitic analysis of the copula. The second kind of argument is a new argument from the ‘Echo Contrastive Construction’.

Cho and Sells (1995) argue that two phonological rules, Palatalization and Coda Neutralization, apply in strictly lexical rather than phonological domains, and hence can distinguish affixes from clitics. The rule of Palatalization palatalizes a dental preceding the high front vowel *i*, applying only in derived lexical environments which correspond to a single lexical word. So, while the rule applies in examples like (30), it does not in (31), as these are either underived forms or derived forms which correspond to more than one lexical word (i.e., a compound):

(29) **Lexical Palatalization**

$t, t^h \rightarrow c, c^h / _ i$

(30) Palatalization applies:

- a. $kat^h+i <kath+i> \rightarrow kac^hi$
same-ADV ‘together’
- b. $hæ\ tot+i <hay\ tot+i> \rightarrow hætøji$
sun rise-NOMIN ‘sun rising’
- c. $pat^h+i <path+i> \rightarrow pac^hi$
field-NOM
- d. $pat^h+i+ta <path+i+ta> \rightarrow [pac^hida]$
field-COP-DECL
- e. $mat+i+ta <mat+i+ta> \rightarrow majida$
eldest-COP-DECL

(31) Palatalization does not apply:

- a. $ət̪i <eti> \rightarrow ədi$ (*əji)
‘where’
- b. $canti <canti> \rightarrow candi$ (*canji)
‘grass’
- c. $pat^h+ilɑŋ <path+ilɑŋ> \rightarrow pad\ irɑŋ$ (*pajirɑŋ), (pannirɑŋ)
field ridge ‘the ridge of a field’

The fact that there is no Palatalization inside underived words, nor between the two members of a compound (as in *pat^h+ilɑŋ*) indicates that the domain of the rule is not the Phonological Word, but a smaller sublexical constituent. Crucially, the copula triggers Palatalization in examples like *majida* (< *mat-i-ta*). If it were a clitic, rules sensitive to morphological information should not apply (they would not have access to the requisite information to apply).

The second lexical rule, Coda Neutralization, interacts closely with lexical syllabification and shows that the domain of syllabification is the sublexical constituent, rather than the phonological word. There is a set of coda rules which neutralize a continuant to a stop, a palatal to a dental, and a laryngeal consonant

to a plain stop only in the syllable coda position. However, if the continuant in the stem is syllabified as the onset of the following suffix, it escapes the application of Coda Neutralization, as in seen in the examples *usim*, *osil* and *osita* in (32). Lexical syllabification takes as its domain the stem and the suffixes (either derivational or inflectional) but never covers two separate lexical words. This is why there is obligatory Neutralization in compounding (*k'odirim*) (and *pad iran* in (31)).

(32) Coda Neutralization does not apply:

- a. us+iim <us+um> → usim (*udiim)
smile-NOM 'smile'
- b. os+il <os+ul> → osil
clothes-ACC
- c. os+i+ta <os+i+ta> → osida
clothes-COP-DECL

(33) Coda Neutralization applies:

- k'oc^h+ilim <kkoch+ilum> → k'ot+ilim → k'odirim/
k'onnim
flower-name 'the name of a flower'

Both of these phonological processes show that the copula is not just part of the same Phonological Word as its N host, but part of the same Lexical Word, which is inconsistent with the clitic analysis.

5. Lexical Status III: The Echo Contrastive Construction

The Echo Contrastive Construction (ECC) allows the doubling of Vs but none of their phrasal arguments and adjuncts, and sets up a negative implicature in the interpretation of the whole sentence (see Nishiyama and Cho (1998), Choi (2003), Kim (2002), Jo (2003)). The domain of the Korean ECC construction is in fact any syntactic V^0 (see Cho et al. (2003)):

- (34) a. chelswu-ka seoul-ul pangmwun ha-ki-nun hay-ss-ta
chelswu-NOM Seoul-ACC visit do-NOMIN-FOC do-PAST-DECL
'Chelswu did visit Seoul (but ...).'
- b. chelswu-ka seoul-ul pangmwun ha-ki-nun pangmwun hay-ss-ta
chelswu-NOM Seoul-ACC visit do-NOMIN-FOC visit do-PAST-DECL
'Chelswu did visit Seoul (but ...).'

(34)b is possible as an ECC as *pangmwun ha-ta* is a syntactically complex V^0 – a V with a preceding VN complement. Just the final V alone can be copied, for it is also a V^0 , as in (34)a.

Crucially, the positive copula behaves differently, and the only grammatical form of an ECC with the positive copula also involves doubling the preceding N, as in (35) (see Oh (1991), Kim and Chung (2002)). Under the clitic analysis of the copula, the copula and its complement are never a syntactic constituent (see (3) above); however, the ECC treats N+Copula as a syntactic constituent, and so *pangmwun ha-ki-nun hay-ss-ta* above contrasts the impossible **haksayng-i-ki-nun-i-ta* version of (35). This string ought to be fully

acceptable under the clitic analysis of the copula, with the first clitic *-i-* hosted at PF by *haksayng* and the second clitic *-i-* hosted at PF by *-nun*.⁷

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

Needless to say, the ECC facts with the negative copula are different: just the *ani-* part can be doubled, as in (36)a. And in fact while the doubling of N + negative copula as in (36)b is grammatical, this example does not have the ‘negative implicature’ interpretation typical of the ECC, but rather has a VP-topic interpretation:

- (36) a. papo-ka ani-ki-nun ani-ta
 fool-NOM NCOP-NMLZ-TOP NCOP-DECL
 ‘He’s not a fool (but he’s not so smart).’
 b. papo-ka ani-ki-nun papo-ka ani-ta
 fool-NOM NCOP-NMLZ-TOP fool-NOM NCOP-DECL
 ‘It is true that he’s not a fool.’

In summary, the ECC treats the negative copula *ani-ta* as being a single syntactic V^0 which takes a phrasal complement, while the only V^0 the ECC has access to with the positive copula is the V^0 formed by an N hosting *-i-*, which is effectively a ‘verbalizing’ or ‘copularizing’ suffix, as we now explain.

6. Formal Analysis with Lexical Sharing

We believe that the facts in the previous sections call out for an analysis of the positive copula that is grounded in the lexicality of the N+copula complex. Here we offer such an analysis, using lexical sharing, designed by Wescoat (2002) to handle numerous phenomena across languages in which a word that is demonstrably an output of the lexicon – and not a postsyntactic amalgam – exhibits complex syntactic behavior suggesting that it corresponds to two or more lexical-category constituents in phrase structure. In 6.1. we describe lexical sharing, as rendered within the overarching theory of HPSG (Pollard and Sag 1994, Ginzburg and Sag 2000); then we discuss its application to the foregoing Korean facts in 6.2.

As an example of an application of lexical sharing to a well-known phenomenon, consider English pronoun-auxiliary contractions like *I’ll*. Di Sciullo and Williams (1987) regard *I’ll* as a post-syntactic amalgam, or *phonological word*. However, Spencer (1991) points out that some monosyllabic pronoun-auxiliary contractions show signs of lexical derivation. For instance, it is typical for lexical processes to exhibit a high degree of selectivity in the stems to which they apply; when anglicized, the Japanese given name *Ai* is a homophone of *I*, yet as (37) shows, only the pronoun allows the monosyllabic contraction [ajl].

- (37) a. I’ll [ajl] help.
 b. Ai’ll [aj] / *ajl] help.

Additionally, monosyllabic *I’ll* allows for idiosyncratic stem allomorphy, another lexical characteristic; it admits the pronunciation [al], in which the glide of the pronoun *I* [aj] is lost. Despite the manifest lexicality of monosyllabic *I’ll*, Di Sciullo and Williams are correct in assuming that it is not syntactically simplex. Note that the second conjunct in (38) is a VP which obviously takes a first-person subject.

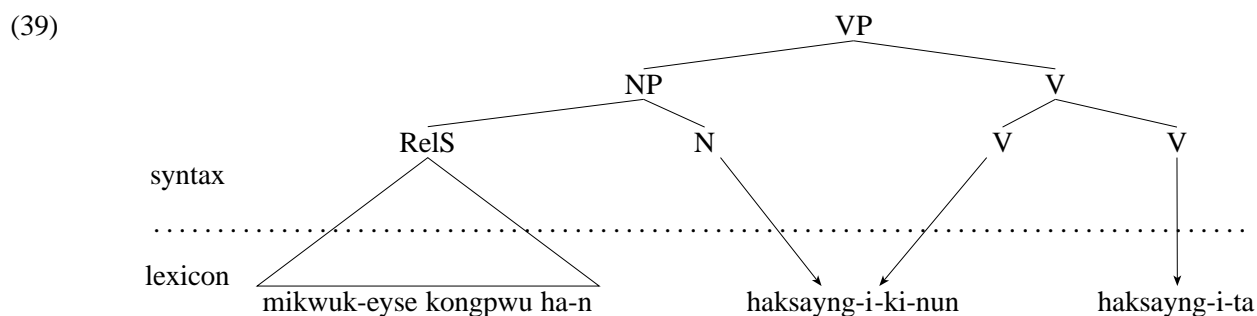
⁷Under the clitic analysis of the copula, the copula and its complement are never a syntactic constituent, so it must be that ECC would have to apply only to phonological constituents (as suggested in the analysis of Jo (2003)). However, we believe that this makes incorrect predictions in other cases (in particular, it predicts **John-i Tom-ul an manna-ki-nun manna-ss-ta* to be grammatical).

(38) I'll [ajl / al] do whatever's necessary and am confident of success.

The simplest way to have parallelism of syntactic categories across conjuncts and to provide the second conjunct with an appropriate subject is to assume that *I'll* instantiates both a first-person pronoun and an auxiliary verb in the phrase structure; the verb forms part of the first conjunct of the coordinate structure, while the pronoun lies outside. The notion of lexical sharing outlined above provides for the combination of lexical unity and syntactic complexity displayed by monosyllabic *I'll*, which other theories would deem paradoxical.

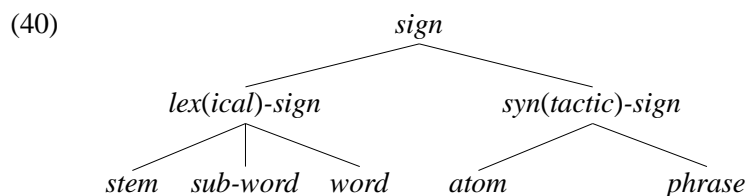
6.1. Lexical Sharing and HPSG

Wescoat (2002) assumes 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 into a *lexical exponent*, i.e. a word which is said to *instantiate* the atomic constituent. The basic idea of lexical sharing is then that two or more atomic constituents may 'share' the same exponent, or equivalently, that a single word may instantiate multiple atomic constituents. This scheme provides a straightforward model of words that appear to straddle a phrase boundary. Diagrammatically, we distinguish the exponence relation from the usual dominance relation by employing arrows in place of simple lines, as illustrated in (39), which outlines our analysis of the lexicity of *haysayng-i-* in (35)a.



Crucially for our purposes, it is a theorem of Wescoat's axiomatization of lexical sharing that a single word can be the exponent of multiple atom only if those atoms are adjacent.

Lexical sharing may be simply 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, for our purposes here, is shown in (40).



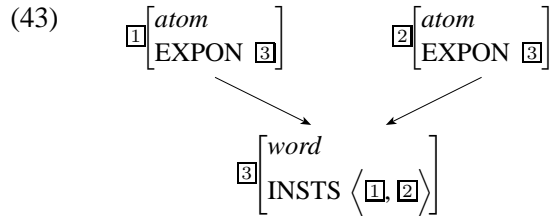
The type (in *italics*) of an AVM determines what attributes (in SMALL CAPS) and what types of values the AVM may contain. The principal new type declarations are given in (41).

- (41) a.
$$\left[\begin{array}{l} \textit{lex-sign} \\ \text{PHON(OLOGY)} \textit{ list(form)} \\ \text{INST(ANTIATE)S} \textit{ non-empty-list(atom)} \end{array} \right]$$
- b.
$$\left[\begin{array}{l} \textit{syn-sign} \\ \text{SYNSEM} \textit{ synsem} \end{array} \right]$$
- c.
$$\left[\begin{array}{l} \textit{atom} \\ \text{EXPON(ENT)} \textit{ word} \\ \text{ARG(UMENT)-ST(RUCTURE)} \textit{ list(synsem)} \end{array} \right]$$
- d.
$$\left[\begin{array}{l} \textit{phrase} \\ \text{D(AUGH)T(E)RS} \textit{ non-empty-list(syn-sign)} \end{array} \right]$$

The new attributes EXPON and INSTS in (41) implement lexical sharing: each *atom* is linked via EXPON to a lexical exponent of type *word*; additionally, every *word* contains, as the value of INSTS, an ordered list enumerating each *atom* that the *word* instantiates. We ensure reciprocal linkage between *word* and *atom* with the constraints in (42).

- (42) a.
$$\textit{atom} \Rightarrow \boxed{1} \left[\text{EXPON} \left[\begin{array}{l} \textit{word} \\ \text{INSTS} \langle \dots, \boxed{1}, \dots \rangle \end{array} \right] \right]$$
- b.
$$\textit{word} \Rightarrow \boxed{1} \left[\text{INSTS} \left\langle \left[\begin{array}{l} \textit{atom} \\ \text{EXPON} \boxed{1} \end{array} \right], \dots, \left[\begin{array}{l} \textit{atom} \\ \text{EXPON} \boxed{1} \end{array} \right] \right\rangle \right]$$

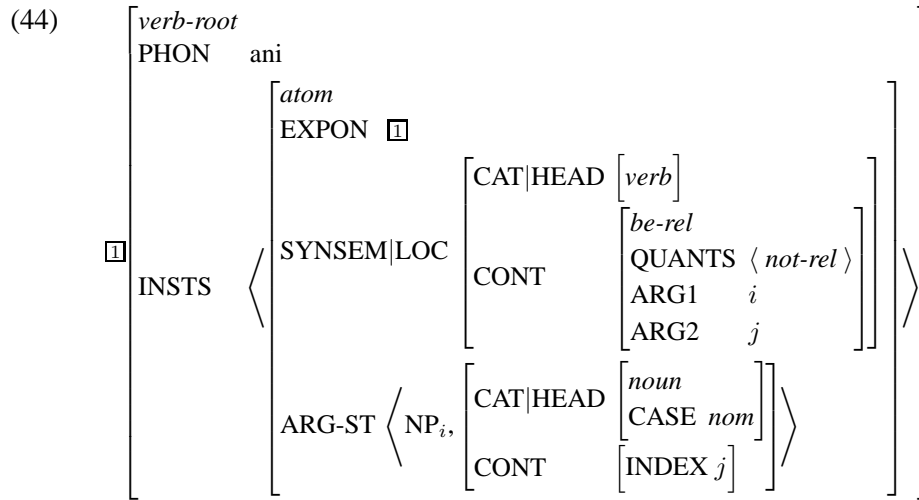
The effect of (42) is illustrated by the schematization in (43) of an instance of lexical sharing (compare with the *haksayng-i-ki-nun* part of (39)).



The tags $\boxed{1}$, $\boxed{2}$, and $\boxed{3}$ reveal the interpenetration of the AVMs which they index: both *atom* $\boxed{1}$ and *atom* $\boxed{2}$ have the same *word* $\boxed{3}$ as value of EXPON, giving rise to lexical sharing; moreover, *word* $\boxed{3}$ contains both *atom* $\boxed{1}$ and *atom* $\boxed{2}$ in its INSTS list, thereby enabling the *word* to determine individually the syntactic features of each *atom*. Of course, when there is no actual lexical ‘sharing’, i.e. when a *word* is exponent of a single *atom*, the INSTS list is simply of length one.

6.2. Copularization

The analysis of *-i-* will be easier to see if we start with the negative copula first. The entry for this form is shown in (44) (we assume a sort *verb-root* for the basic representation of verb roots in Korean):

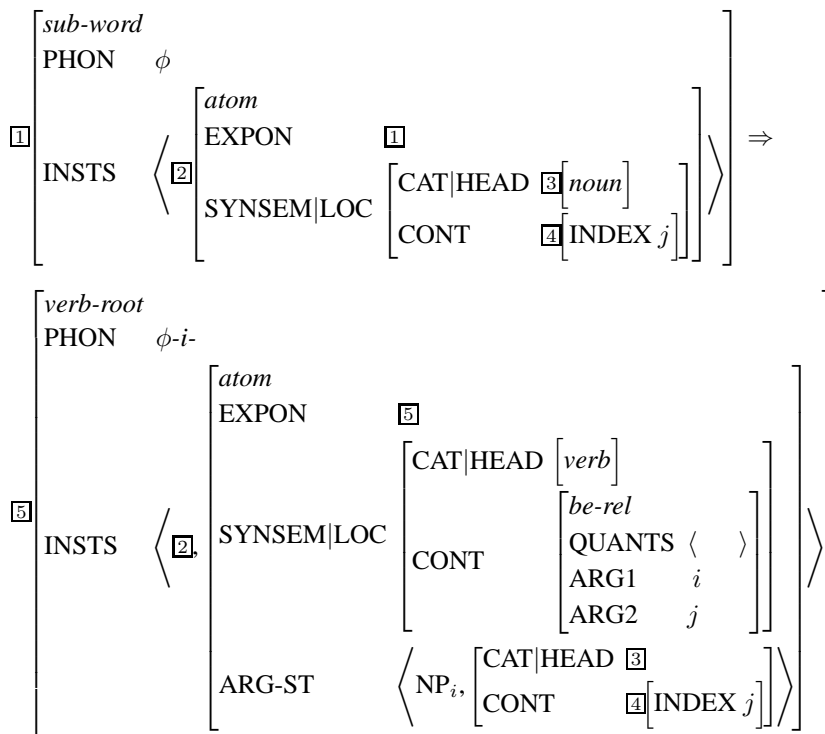


This is the entry of a two-place predicate which expresses negation of the *be-rel* between those arguments. We adopt the analysis of negation in Ginzburg and Sag (2000), where *neg-rel* appears on the QUANTS list. The first argument is the subject, and the second argument is a nominative-marked complement.

Such an entry would also be appropriate for a positive copula, without the *neg-rel*. The different uses of the full-fledged copula mentioned in section 2 would involve related entries that differ in terms of information-structure properties and also possibly in terms of subtly different sub-types of *be-rel*. The fragment use of the copula inherits the value of the semantic relation from the context.

Now to the positive copula *-i*, which is the result of a rule which takes a noun *sub-word* as input, and which returns a *verb-root*. That new lexical item can instantiate two syntactic atoms, the first of which is the noun that was input to the rule, and the second is a two-place predicate which expresses the *be-rel*:

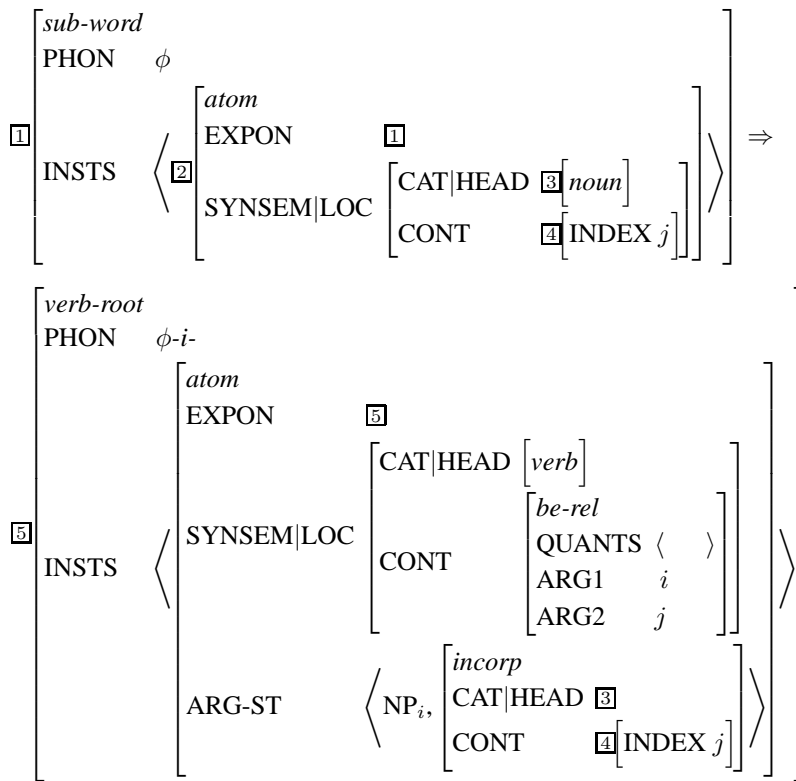
(45) **Copularization Rule**



This lexical rule applies to an N sub-word and creates a form of type *verb-root*. That new 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. Hence, this is appropriate for the form *haksayng-i-ki-nun* in (39). 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 then a subject NP.

In contrast, the final word *haksayng-i-ta* in (39) only instantiates one atom, V. Effectively, *haksayng* is truly incorporated into the verb. The rule above therefore has a different output, as shown in (46). In this analysis we follow the account of incorporation in Malouf (1999), which builds on the HPSG distinction between *canonical* and *non-canonical* SYNSEMs (see e.g., Ginzburg and Sag (2000, 40)). The second element on the ARG-ST list is of a *non-canonical* type, namely *incorp(orated)*, which means that although it is present in the lexical entry and in the semantics, it does not correspond to an overt dependent of the verb in the syntax.

(46) **Copularization Rule, Incorporating Version**



7. Conclusion

Under the Lexical Sharing account developed above, *haksayng-i-* is formed in the lexicon and may then be input to any processes of verbal affixation. As it is a lexical unit, it necessarily has the properties enumerated in sections 3–5 above, properties which are either accidental or uncapturable on accounts that do not recognize the lexicality of this form. The Lexical Sharing approach nevertheless allows this single word to instantiate the two atoms N and V in (3) above, showing how a morphologically unitary form can participate in complex syntactic behavior.

References

- Cho, Sae-Youn, Jong-Bok Kim, and Peter Sells. 2003. Contrastive verb constructions in Korean. Paper presented at ISOKL-10, Harvard.
- Cho, Young-mee Yu, and Peter Sells. 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*, Vol. 11. CSLI, Stanford Linguistics Association, 457–470.
- Di Sciullo, Anna-Maria, and Edwin Williams. 1987. *On the Definition of Word*. Cambridge, MIT Press.
- Ginzburg, Jonathan, and Ivan Sag. 2000. *Interrogative Investigations*. CSLI Publications.
- Han, Hak-Sung. 1996. Syntactic affixation and the structure of Korean. In Hee-Don Ahn, Myung-Yoon Kang, Yong-Suck Kim, and Sookhee Lee (eds.), *Morphosyntax in Generative Grammar (Proceedings of 1996 Seoul International Conference on Generative Grammar)*. Seoul, The Korean Generative Grammar Circle, Hankwuk Publishing Co., 23–37.
- Heycock, Caroline, and Anthony Kroch. 2002. Topic, focus, and syntactic representations. In Line Mikkelsen and Chris Potts (eds.), *Proceedings of WCCFL 21*. Somerville, Mass., Cascadilla Press, 145–162.
- Jo, Jung-Min. 2003. Variation in predicate cleft constructions in Korean: Epiphenomena at the Syntax-PF interface. Paper presented at ISOKL-10, Harvard.
- Kim, Jong-Bok, and Chan Chung. 2002. Korean copula constructions: A construction and linearization perspective. *Ene* 27, 171–193.
- Kim, Youngsun. 2002. The syntax of Korean verbal focus constructions. In *Proceedings of the 2002 LSK International Summer Conference*. The Linguistic Society of Korea, 333–344.
- Malouf, Robert. 1999. West Greenlandic noun incorporation in a monohierarchical theory of grammar. In G. Webelhuth, J.-P. Koenig, and A. Kathol (eds.), *Lexical and Constructional Aspects of Linguistic Explanation*. Stanford, CSLI Publications, 47–62.
- Mikkelsen, Line. 2002. Specification is not inverted predication. In Masako Hirotani (ed.), *Proceedings of NELS 32*. Amherst, GLSA, 403–422.
- Nishiyama, Kunio, and Eun Cho. 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*, Vol. 7. CSLI, Stanford Linguistics Association, 463–479.
- Oh, Mira. 1991. The Korean copula and palatalization. *Language Research* 27, 701–724.
- Pollard, Carl, and Ivan Sag. 1994. *Head-Driven Phrase Structure Grammar*. Chicago, University of Chicago Press and Stanford, CSLI Publications.
- 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*, Vol. 7. Dept. of Linguistics, Harvard University, 488–503.

- Spencer, Andrew. 1991. *Morphological Theory*. Oxford, Blackwell.
- Stump, Gregory. 2001. *Inflectional Morphology: A Paradigm Structure Approach*. Cambridge, Cambridge University Press.
- Wescoat, Michael T. 2002. *On Lexical Sharing*. Doctoral dissertation, Stanford University.
- Yang, In-Seok. 1972. *Korean Syntax: Case Marking, Delimiters, Complementation and Relativization*. Seoul, Paek Hap Sa.
- Yoon, James Hye-Suk. 1995. Nominal, verbal, and cross-categorical affixation in Korean. *Journal of East Asian Linguistics* 4, 325–356.
- Yoon, James Hye-Suk. 2002. Non-morphological determination of nominal affix ordering in Korean. In Lorie Heggie and Francisco Ordonez (eds.), *Theoretical Perspectives on Clitic and Affix Combinations*. Amsterdam/Philadelphia, John Benjamins, to appear.
- 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*, Vol. 11. CSLI, Stanford Linguistics Association, 34–49.