

Bracketing Paradoxes in Natural Language Processing*

Jay Jeong-Ryeol Kim

This paper describes two different types of suffixes in Korean depending on whether one class of suffixes causes bracketing paradoxes or not. It explains how the ECS bidirectional Korean-English translation system deals with the bracketing paradoxes triggered by a group of suffixes.

1. Introduction

In so-called 'bracketing paradoxes' (Kiparsky (1983), Pesetsky (1979, 1985), Williams (1981)) it is necessary to assign two incompatible constituent structures, morpho(phono)logical and semantico-syntactic structures, to a single expression. For example,

- (1) a. *Tom's a linguist.* (Spencer (1988))
- b. *unhappier* (Pesetsky (1985))

In (1a) *Tom's* behaves as a single word morpho(phono)logically, but *'s a linguist* forms a constituent syntactically as shown in (2) :

- (2) [[Tom's] a linguist] morpho(phono)logical
- [Tom['s a linguist]] syntactic

In (1b) morpho(phono)logically *happier* forms a primary constituent, but semantically *unhappy* forms a primary constituent as shown in (3) :

- (3) [un[happier]] morpho(phono)logical
- [[unhappy]er] semantic

*I wish to thank Joseph Pentheroudakis, Dan Higginbothom, One-soon Her and Deborah Masterson for their comments on earlier versions of this paper. The remaining errors, of course, are mine.

In Korean we can find various kinds of constructions which trigger bracketing paradoxes : copula suffix *-i-*, coordinator *-kwa*, postpositions *-eyse* (location), *-losse* (means), *-lose* (qualification) and nominalizer *-ki*.

The purpose of this paper is (1) to show that there are suffixes that create bracketing paradoxes and others that do not; and (2) to suggest a solution to such bracketing paradoxes in lexical and/or syntactic processing.

2. Morpho(phono)logical Bracketing

The following evidence shows that the relationship between base forms and the suffixes triggering bracketing paradoxes is cohesive as a word.

1) Case marking is not allowed inside word boundaries in Korean (Kim (1990), O'Grady (1990), Shibatani et al. (1988)). The noun plus a suffix, the copula *-i-* for example, ban the intervention of a case marker between them (Song (1988)) :

- (4) *ku kyoswu -ka hakca (*-ka) i-ta*
 the professor-NOM scholar -NOM be -DECL
 'The professor is a scholar.'

2) A parenthetical expression which would normally be acceptable between two syntactic constituents is not allowed inside word boundaries (Kim (1990), Shibatani et al. (1988))

- (5) *ku kyoswu -ka hakca (*nay chwuchuk-ey) i-ta*
 the professor-NOM scholar (my guess-LOC) be-DECL
 'The professor is, in my guess, a scholar.'

3) An element of a morpho(phono)logical constituent cannot be scrambled with other elements in a sentence out of the boundary of word (O'Grady (1990))

- (6) **hakca ku kyoswu-ka i -ta*
 scholar the professor be-DECL
 'The professor is a scholar.'

4) Morpho(phono)logical constituents show the following phonological characteristics which do not apply across the word boundary :

i) We can observe that the following lexical level phonological contraction applies to newly formed words : intervocalic glidification in (7) and syllable deletion influenced by speech style in (8). Note that such contraction is generally not acceptable across word boundary (or boundaries).

- | | |
|--|---|
| <p>(7) $i \rightarrow y / V + _ + V$
 / emeni + i + ess + / \rightarrow
 mother-COPULA-PST
 [emeni + y + ess +]
 ‘was a mother’</p> | <p>$*i \rightarrow y / V \# - + V$
 / himepsi # i + ess + / \rightarrow
 weakly carry[on head]-PST
 $*[himepsi \# y + ess]$
 ‘carry on head weakly’</p> |
| <p>(8) a. $eyse \rightarrow se / \text{noun} + _$
 (casual speech)
 / Seoul + eyse / \rightarrow [Seoul + se]
 Seoul LOC
 ‘in Seoul’</p> | |
| <p>b. $i \rightarrow O / V + _$
 (casual speech)
 / emeni + i + ta / \rightarrow
 mother-COPULA-DECL
 [emeni + ta]
 ‘is a mother’</p> | <p>$*i \rightarrow O / V \# _$
 (casual speech)
 / himepsi # i + ta / \rightarrow
 weakly carry[on head]-DECL
 $*[himepsi \# ta]$
 ‘carry on head weakly’</p> |

ii) i insertion occurs when the l-initial suffix is preceded by a consonantal ending as in (9).

- | | |
|---|---|
| <p>(9) $O \rightarrow i / C + _ 1$
 / chayk + losse / \rightarrow
 book -P
 [chayk + i losse]
 ‘by a book’</p> | <p>$*O \rightarrow i / C \# _ 1$
 / posek # latio / \rightarrow
 jewelry radio
 $*[posek \# ilatio]$
 ‘jewelry radio’</p> |
|---|---|

iii) A consonant is deleted when the coordinator -kwa, for example, is preceded by a vowel ending as in (10).

- | | |
|---|--|
| <p>(10) $C \rightarrow O / V + _$
 / emeni -kwa / \rightarrow
 mother-CORD
 [emeni - wa]
 ‘and mother’</p> | <p>$*C \rightarrow O / V \# _$
 / cenki # kwahak / \rightarrow
 electricity science
 $*[senki \# wahak]$
 ‘science of electricity’</p> |
|---|--|

iv) The tensing of a stop consonant when preceded by another stop is obligatory across a morpheme boundary, though it is optional across a word boundary.

- (11) stop → tensed stop / stop + _(obligatory)
/ stop # _(optional)

/ppop +ki/ → [ppop+kki], but *[ppop+ki]

choose-NMLZR

'choosing'

/penyek # kiswul/ → [penyek # kkiswul] or [penyek#kiswul]

'translation skill'

v) Another obvious clue to the more cohesive relationship between base forms and suffixes is that all the suffixes discussed above are bound morphemes and bound morphemes are attached orthographically to the previous base forms. Thus, the space, which often serves as a lexical boundary for language processing by computer, is not inserted between a base form and a suffix, while it is used between words.

- (12) say # chayk + i + ta →
new book- COPULA-DECL
say chaykita
'is a new book.'

Based on the above morpho(phono)logical evidence the following bracketing is appropriate for the expressions :

- | | |
|---------------------------|------------------------------------|
| (13) a. copula suffix -i- | b. coordinator -kwa |
| <i>say chayk-i -ta</i> | <i>say chayk-kwa...</i> |
| new book -COPULA-DECL | new book -CORD |
| 'is a new book' | 'a new book and ...' |
| [say [chayk-i-ta]] | [say [chayk-kwa]] |
| c. postposition | d. nominalizer -ki |
| <i>say chayk-eyse</i> | <i>aki- ka acik uyca-ey anc-ki</i> |
| new book -LOC | baby-NOM yet chair-LOC sit-NMLZR |
| 'in a new book' | -ka <i>elyep- ta</i> |
| [say [chayk-eyse]] | -NOM difficult-DECL |

‘It is still difficult for the baby to sit
on a chair’
[[aki-ka acik uyca-ey [anc-ki-
ka]]elyep-ta]

3. Syntactic Bracketing

However, there is syntactic evidence that the bracketing should be done in another way.

i) The incorporated nouns to the suffixes above can have an external modifier.

- (14) a. *say chayk- i- ta*
new book -be-DECL
‘[It] is a new book’
b. *say chayk-eyse*
new book -LOCT
‘in a new book’
c. *say chayk-ulosse*
new book -NST
‘with a new book’

In contrast with the examples in (14), there are incorporated nouns to the following suffixes that do not allow an external modifier.

- (15) a. **say inkan -cek*
new human-SFX
‘new humane’
b. **say elwuk-ci- ta*
new blot -SFX-DECL
‘blot new’
c. **say caknan-sulep-ta*
new play -SFX -DECL
‘be new playful’

Instead, these nouns allow only adverbial modifiers in the preceding position. This, again, is very different from the previous set of suffixes which

behave in exactly the opposite manner, not allowing adverbial modifiers.

- (16)a. *acwu inkan- cek/* -i -ta*
 very human-SFX/ -be-DECL
 ‘...very humane/be a very human’
- b. *simhakey elwuk-ci-ta/* -i -ta*
 badly blot -SFX-DECL/-be-DECL
 ‘...be blotted badly/badly be a blot’
- c. *acwu calang-sulep-ta /*-i -ta*
 very pride -SFX -DECL/-be-DECL
 ‘...very proud/be a very pride’

2) The nouns incorporated to the suffixes in (14) can have an external possessor or a specifier (Kim (1990)).

- (17) a. *Swuni-uy /ku chayk-i -ta*
 Swuni-GEN/the book-be-DECL
 ‘[It] is Swuni’s/the book.’
- b. *Swuni-uy /ku chayk-eyse*
 Swuni-GEN/ the book -in
 ‘In Swuni’s / the book.’
- c. *Swuni-uy /ku chayk-ulosse*
 Swuni-GEN/the book -NST
 ‘With Swuni’s/the book.’

3) Gapping may occur in a coordinated construction :

- (18) a. *ikes-un chayk, cekes-un kongchayk-i -ta*
 this-TOP book that -TOP notebook -be-DECL
 ‘This is a book, that is a notebook.’
- b. *ikes-un chayk, cekes-un kongchayk-eyse po -ass -ta*
 this-TOP book, that-TOP notebook -LOC see-PAST-DECL
 ‘[I] saw this in a book, that in a notebook.’

4) The incorporated elements may carry an inflectional ending such as number or delimiter for nouns and honorific marker for verbs (Kim (1990))

- (19) a. *Wuli-nun haksayng-tul-i- ta*
 we -TOP student -PL-be-DECL
 ‘We are students.’

- b. *yekiey haksayng-tul -ppwun-i -ta*
 here student -PL -only -be- DECL
 ‘Here are only students.’
- c. *emenim-i o -si -ki -ka eleyep -ta*
 mother-NOM come-HON NMLZR-NOM difficult-DECL
 ‘[My] mother’s coming is difficult.’

5) The incorporated elements plus those suffixes, copula *-i-*, postposition *-eyse* and nominalizer *-ki*, are highly productive and regular phenomena. In fact, it is hard for any native speaker to come up with a word of the appropriate category which fails to participate in incorporation.

6) The external arguments of verbs incorporated to the suffixes in (14) may be present.

- (20) *aki -ka acik uyca -ey anc-ki -ka eleyep -ta*
 baby -NOM yet chair -LOC sit -NMLZR-NOM difficult-DECL
 ‘It is still difficult for the baby to sit on a chair’

On the other hand, the nominalizer *-(u)m* does not allow an incorporated verb to have its external arguments.

- (21) **aki -ka acik uyca -ey anc-um -i eleyep -ta*
 baby-NOM yet chair -LOC sit -NMLZR-NOM difficult-DECL
 ‘It is still difficult for the baby to sit on a chair’

However, the nominalizer *(u)m* may have an external possessor (22a), nominal modifier (22b) and/or specifier (22c). This is, again, very different from the nominalizer *-ki*, which is shown as ungrammatical in the same environment.

- (22) a. *haksayngtul-uy ssawu -m /*ki*
 students -GEN fight -NMLZR/NMLZR
 ‘students fight’
- b. *say ssawu-m /*ki*
 new fight-NMLZR/NMLZR
 ‘new fight’
- c. *ku ssawu-m /*ki*
 the fight-NMLZR/NMLZR
 ‘the fight’

Also, the nominalizer $-(u)m$ can further undergo a word formation rule such as $N + cil \rightarrow N$, but the nominalizer $-ki$ cannot.

- (23) a. $N + cil$ 'action' $\rightarrow N$
 as in *totwuk-cil* 'stealing'
- tali-m/*ki-cil* 'ironing'
*ttak-um/*ki-cil* 'wiping'
*kkak-um/*ki-cil* 'sharpening'
*sokk-um/*ki-cil* 'thinning'
*pak-um/*ki-cil* 'sewing'

The syntactic evidence discussed above shows that these incorporated base forms to the suffixes ($-i-$, $-eyse$, $-losse$, $-kwa$ and $-ki$) maintain the same syntactic properties and act as a head element at the syntactic level: being able to take an external modifier, possessor and/or specifier and argument(s). Thus, the morpho(phono)logical bracketings of (13) are contrasted with the syntactic bracketings of the same sentences in (24):

- (24) a. copula suffix $-i-$
say chayk -i -ta
 new book -COPULA-DECL
 'is a new book'
 [say [chayk-i-ta]] morpho(phono)logical
 [[say chayk]-i-ta] syntactic
- b. coordinator $-kwa$
say chayk-kwa...
 new book -CORD
 'a new book and...'
 [say [chayk-kwa]] morpho(phono)logical
 [[say chayk]-kwa] syntactic
- c. postposition
say chayk-eyse
 new book -LOC
 'in a new book'
 [say [chayk-eyse]] morpho(phono)logical
 [[say chayk]-eyse] syntactic
- d. nominalizer $-ki$
aki -ka acik uyca -ey anc-ki -ka elyep -ta

baby-NOM yet chair -LOC sit -NMLZR-NOM difficult-DECL
 'It is difficult that the baby sits on a chair yet'
 [[aki-ka acik uyca-ey[anc-ki-ka]]elyep-ta] morpho(phono)-
 logical
 [[[aki-ka acik uyca-ey anc]-ki-ka]elyep-ta] syntactic

The suffixes showing bracketing paradoxes in (24) are in contrast with the suffixes with no apparent contradiction between morpho(phono)logical constituency and syntactic constituency as in (25):

- (25) a. **say/satlopkey inkan-cek*
 new/newly human-SFX
 'new/newly humane'
 [saylopkey [inkan-cek]] morpho(phono)logical and syntactic
- b. **say/saylopkey elwuk -ci -ta*
 new/newly blot -SFX-DECL
 'get a new blot'
 [saylopkey [elwuk-ci-ta]] morpho(phono)logical and syntactic
- c. **say/saylopkey calang -sulep-ta*
 new/newly pride -SFX -DECL
 'be new/newly proud'
 [saylopkey [calang-sulep-ta]] morpho(phono)logical and syntactic
- d. *say / *saylopkey ssawu-m. -i -sicktoy-ess -ta*¹⁾
 new/newly fight -NMLZR-NOM start -PST-DECL
 'A new/newly fight is started'
 [new [ssawu-m]] morpho(phono)logical and syntactic

¹⁾ One has to note that the sentence is ungrammatical in the interpretation in which the adverbial modifier *saylopkey* 'newly' modifies the verb *ssawu*-fight'. I also have to emphasize that the suffix *-(u)m* is a homophone, one version is a stylistic marker and the other a nominalizer. The stylistic marker *-(u)m* is used in a sentential final position effecting a shortened form especially in military documents as in the following example:

samangka-ka 25 myeng palsaynghayss-um
 dead -NOM 25 people occur -SFM
 '25 people dead'

This use of *-(u)m* is excluded from our discussion.

4. Phrasal Suffixes vs. Lexical Suffixes

As discussed in the previous section there are two sets of suffixes: the first set defers the category change of the base form the suffixes attach to until the base form expands to its maximal projection as a head: *-i-*, *-eyse*, *-lose*, *-losse* and *-ki*. The other set changes the category of a base form immediately after the attachment: *-cek*, *-ci-*, *-sulep-* and *-(u)m*. I will call the former as “phrasal suffixes” and the latter as “lexical suffixes”.

Following Selkirk(1982) with slight modification, I assume that lexical entries representing suffixes will include information on a) category, b) subcategorization frames: both morpho(phono)logical and syntactic if necessary, c) semantic functions and d) phonological representation. As an illustration, the entries for *-i-*, *-eyse*, *-sulep-* and *-m* are represented in the following manner:

(26) <i>-i-</i> : Category V^{af}	<i>-eyse</i> : Category P^{af}
Subcategorization	Subcategorization
[N-]morpho(phono)logical	[N-] morpho(phono)logical
[NP-]syntactic	[NP-] syntactic
Semantic function	Semantic function
Phonological representation	Phonological representation
/i/	/eyse/
<i>-sulep-</i> : Category V^{af}	<i>-m</i> : Category N^{af}
Subcategorization	Subcategorization
[N-]	[V-]
Semantic function	Semantic function
Phonological representation	Phonological representation
/sulep/	/m/

The phrasal suffixes generate bracketing paradoxes owing to the contradicting bracketing between morpho(phono)logical and syntactic constituents. (27) shows the word formation rule (c) and phrase structure rules (a, b) for words containing phrasal suffixes. Note that the same category X represents both the maximal projection in (27a), XP, and the suffixes in the word formation rule (27c), X^{af} .

- (27) a. $XP \rightarrow YP$
 b. $YP \rightarrow \alpha Y$
 c. $Y \rightarrow Y + X^{af}$

where each X and Y is a major category: N, A, V, P and $X \approx Y$. Also, α is a variable over a sequence of category symbols which the grammar allows as sisters of Y .

The rules in (27) permit the concept of “acting” head in which the category Y incorporates all the elements in α by virtue of phrasal requirement in the syntactic subcategorization as in $-i-$: $[NP_-]$. The definition of “acting” head is stated in the following:

- (28) Y is an “acting” head if and only if a category Y combines with a suffix X which has a syntactic subcategorization $[YP_-]$, and the YP is not formed (where $X \approx Y$ and X, Y are major categories: N, V, P or A).

In contrast, the lexical suffixes are represented in the following word formation rule:

- (29) $X \rightarrow Y + X^{af}$
 where X and Y are major categories: N, V, P or A .

Bracketing paradoxes raise interesting problems in natural language processing. In parsing, for example, the morpho(phono)logical level is generally processed first to facilitate the dictionary look-up and build a category projected by the head. By the time the processing reaches the syntactic level, the category may have already been transformed by a morphological suffix and an adjacent word cannot participate in building a higher node. A typical example of this is shown in (4b) where the morpho(phono)logical processing is finished, the word *chayk-i-ta* ‘is a book’ forms a V which is projected by the copula $-i-$ (remember Korean is a head-final language), and thus, the adjective *say* ‘new’ cannot be combined with the N *chayk* ‘book’ which has already been incorporated into the V *chayk-i-ta*. In order to prevent the transformed category from blocking a syntactic expansion, I apply the linguistic description explained above. I will discuss (1) the translation system of Executive Communication Systems (ECS) in which I implemented the linguistic description, (2) more specifically the issues related to bracketing paradoxes.

5. The ECS Translation System

The ECS translation system employs an LFG-based indirect transfer method in its translation.²⁾ The relevant parts of the translation system to this paper can be represented as in the Fig. 1.

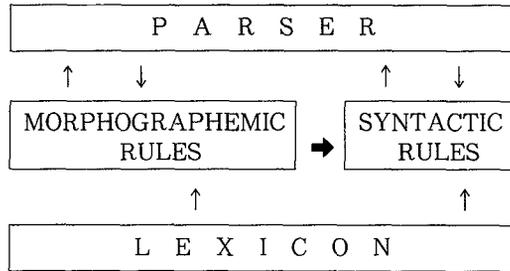


Fig. 1

The parser interacts with morphographic rules and syntactic rules to build complete edges which in turn remain active in the chart for its potential use in the further processing. It uses a breadth-first search and the direction of parsing is flexible either top-down or bottom-up and either left-to-right or right-to-left depending on the specification of linguistic rules.

Both morphographic rules and syntactic rules consist of three components: pattern, rule conditions and action (Her et al. (1989)) :

```

(30) kR_ RULENAME : :
      [ X : α ? < ↓ f1 > =c v1
          pattern Y. ]
      < X f2 > =c < Y f2 >
          rule conditions
      → BUILD (Z)
  
```

²⁾ Refer to Bresnan(1985) for the theory. The motivation in detail behind using LFG as a primary tool of linguistics is well-explained in Her et al. (1989) and Masterson (1990): briefly, it is an established linguistic theory with an efficient processing mechanism: specifically its reversibility between parsing and generation, and the distinction between f- and c-structures which enables the system to cover languages with very different structures. Also refer to Pentheroudakis (1990) for the motivation behind using indirect transfer method.

$$\begin{array}{c} \text{action} \\ \&\& < \uparrow f3 > = < \downarrow f3 > \\ \dots \end{array}$$

where X, Y and Z are grammatical categories, fn is a feature, vn is a value and α represents a grammatical relation.

Pattern also includes a marking of head element represented by “.” and the central element in case of bottom-up rules which will be the starting point of search (left-to-right or right-to-left). Rule conditions is the portion added to the ECS system to increase the expressive power of the grammar. Rule conditions allows specific reference to grammatical categories and thus enables the grammar to block illegal combinations of elements earlier on the parsing process before another category is actually built. Action is the component that builds another category in which function unification occurs. Action could include a bottom-up rule to execute on the next stage based on a currently built category.

Lexicon contains feature inheritance frameworks of a grapheme, a category, information about allomorphic selection, a list of bottom-up rules to execute and functional structure. For example:

(31) kW_word : : [CAT X
MPOST [base_ form_name allomorph_name(s)]
RULES [Y \$ rulename1, rulename2... \$]
FS [f1 v1...fn,vn]]

where X, Y are names of category, f1...fn are features and v1...vn are values.

The functional structure (FS) in a lexical entry is a basic cell to form a larger structure incrementally incorporating information supplied by morphographemic rules and by syntactic rules. When an unknown word is handed on to parser, it goes through a listed set of morphographemic rules either until a morphographemic rule in the list succeeds and the base form is found in a dictionary, or until all the morphographemic rules in the list fail and the word is treated an unknown word.

6. Bracketing Paradoxes in ECS System

Assume that an input sentence is *ikes-i hen chayk-i-ta* ‘this is an old book’. After projections of *ikes-i* (this) and *hen* (old) from dictionary an unknown word to dictionary *chayk-i-ta* ‘is a book’ is passed on to the morphographemic rule called “KR_NOUN”:

(32) $N \rightarrow NSTEM \quad \alpha$
 $\uparrow = \downarrow \quad \uparrow = \downarrow$
 $\{ V^{af}, P^{af} \} \quad \beta$
 $\uparrow = \downarrow \quad (\uparrow SNODE) = \downarrow$
 $(\downarrow mSUBCAT) = {}_cN_-$

kR_NOUN : :

[: NSTEM ./ $\uparrow = \downarrow$ (mPLR \uparrow NUMBER=PL)
plural marker

({ : [mCASE]
[mTPC]
[V^{af}(mHON \uparrow SNODE= \downarrow)
honorification marker
(mTNS \uparrow SNODE= \downarrow)
tense marker
(mMDL \uparrow SNODE= \downarrow)
modality marker
(mFML \uparrow SNODE= \downarrow)
formality marker
(mSFM \uparrow SNODE= \downarrow)
sentence final marker]
[P^{af} (mTPC \uparrow SNODE= \downarrow)
topic marker] : }) :]

\rightarrow BUILD(N)
 && EXECUTE (1, \$ kR_N1 \$)

Since Korean is a suffixal language, morphographemic analysis will proceed right-to-left. This rule will succeed by matching *chayk-i-ta* to the pattern NSTEM-V^{af}-mSFM. The entry for copula *-i-* will contain nominal information on the top level and all the verbal information on the SNODE level in its functional structure:

- (33) $\left\{ \begin{array}{l} \text{CASE MINUS} \\ \text{TOPIC MINUS} \\ \text{SNODE [PRED 'i< (SUBJ) (NCOMP) >']} \end{array} \right.$

The sentence final marker *-ta* will supply to the SNODE the information on speech act (SPACT), declarative (DECL).

When the morphographemic analysis is finished, the requested rule “KR-N1” will be executed. The rule “kR-N1” is as follows:

- (34) kR_N1 ::
 [N ↑ NOMAD = ↓ & &
 ↓ CASE = _cMINUS & &
 ↓ SNODE = MINUS
 N ↑ = ↓]
 → BUILD (N)

The rule “kR_N1” collects prenominal noun specifiers which show exceptional behavior from other modifiers or specifiers in scrambling and cohesion with the head noun, and yet are generally characterized by their possible isomorphic translation between Korean and English such as *hakkyo sinmwun* and *school newspaper*.

- (35) a. *ku cenghwakhan hakkyo sinmwun*
 the correct school newspaper
 ‘the correct school newspaper’
 b. *cenghwakhan ku hakkyo sinmwun*
 correct, the school newspaper
 ‘the correct school newspaper’
 c. **cenghwakhan hakkyo ku sinmwun*
 correct school the newspaper
 ‘the correct school newspaper’
 d. **ku hakkyo cenghwakhan sinmwun*
 the school correct newspaper
 ‘the correct school newspaper’

However, this rule will fail since *chayk-i-ta* is not preceded by another noun in the rule pattern (34). After all the bottom-up rules are executed, the sentence building rule is executed:

(36) kR__*SENT ::

[(NP ↓ CASE=_cNOM & & 'ikes-i'
 ↑ SUBJ= ↓ 'this-NOM')
 PP* ↓ ∈ ↑ ADJUNCTS
 VP ↑ = ↓ 'hen chayk-i-ta'
 'is an old book']

→ BUILD (S)

kR__*VP ::

[[(NP ↓ CASE=_cNOM & &
 ↑ NCOMP= ↓)
 (NP ↓ CASE=_cACC & &
 ↑ OBJ= ↓)
 (NP ↓ CASE=_cDAT & &
 ↑ OBJ2= ↓)
 (PP ↓ PCOMP= ↑ & &
 ↓ PCASE= # LOCT, SRC, ... #)
 V ↑ = ↓]
 [NP ↑ NCOMP= ↓ & &
 'hen chayk-i-ta
 ~ (↓ SNODE=MINUS)]]]
 (FOUNDA (NP, 1) ? <V@I PRED> :- NCOMP)
 →
 BUILD (VP) & &
 (CHKLABEL (↑ , NCOMP)
 & & ~ (< ↑ NCOMP SNODE > =MINUS)
 ? (< ↓ NCOMP SNODE
 PRED> = ↑ PRED & &
 < ↓ NCOMP SNODE
 TENSE> = ↑ TENSE & &
 < ↓ NCOMP SNODE
 SPACT> = ↑ SPACT & &
 < ↓ NCOMP SNODE HON> = ↑ HON & &
 < ↓ NCOMP SNODE
 FORMAL> = ↑ FORMAL & &
 < ↓ NCOMP SNODE
 MODAL> = ↑ MODAL))

kR__*NP ::

[< (DET ↑ DEFINITE=PLUS)

(NP ↑ POSS = ↓)
 (S* ↓ ∈ ↑ ADJUNCTS)
 >
 (ADJ* ↓ ∈ ↑ ADJUNCTS)
 `hen 'old'
 N ↑ = ↓
 chayk-i-ta 'is a book']

→

BUILD (NP)

The sentence building rule “kR_*SENT” puts in a stack the request for the categories VP, NP and PP. The VP building rule “kR_*VP” requests a category NP, PP and V as shown by the rule in (36). The NP building rule “kR_*NP” builds the category NP out of ADJ *hen* ‘old’ and N *chayk-i-ta* ‘is a book’, and another NP out of N *ikes-i* ‘this-NOM’. The VP building rule “kR_*VP” builds the category VP out of NP *hen chayk-i-ta* ‘is an old book’. And then the sentence building rule “kR_*SENT” goes ahead and builds the category S based on the categories NP and VP.

The functional biuniqueness condition is met between the predicate argument structure and the functional structure on the sentential level:

(37) [PRED ‘i<(SUBJ), (NCOMP)>’
 [SUBJ [...]
 [NCOMP [...]

The category S is built and the functional structure of S can be shown as in (38):

(38) [PRED ‘i<(SUBJ), (NCOMP)>’
 TENSE PRES
 SPACT DECL
 SUBJ
 [PRED ‘ikes’
 [TOPIC MINUS
 [CASE NOM
 mSUBCAT N_
 NCOMP [PRED ‘chayk’
 [TOPIC MINUS
 [mSUBCAT N_
 [ADJUNCTS {[PRED ‘hen’]}]

7. Conclusion

I have shown how we can deal with bracketing paradoxes in LFG. The core of LFG, the acknowledgement of f-structure, contributes to the resolution of bracketing paradoxes in such a way that the grammar provides a device by which one extracts multiple levels of f-structure from a single derived word.

Also, I have provided linguistic evidence that we have to classify suffixes into two different classes in terms of their behavior of influencing category switching of base form: phrasal suffixes and lexical suffixes. The phrasal suffixes create bracketing paradoxes by allowing the expansion of base form to its maximal projection, while the newly derived words share morpho(phono)logical properties of words.

In natural language processing, I have successfully applied the linguistically motivated idea of resolving bracketing paradoxes to the ECS bidirectional machine translation system between Korean and English.

References

- Bresnan, Joan (1985) (2nd printing; 1st printing 1982) *The Mental Representation of Grammatical Relations*, Cambridge, Mass.: MIT Press.
- Her, One-Soon et al. (1989) 'An LFG-based Machine Translation System,' 1989 *Proceedings of the International Symposium on Chinese Text Processing*, Boca Raton, Florida.
- Kim, Jay Jeong-Ryeol (1990) 'Noun Incorporation of ha-verbs in Korean,' Paper Presented at the 7th International Conference on Korean Linguistics, Osaka, Japan.
- Kiparsky, Paul (1983) 'Word Formation and the Lexicon,' F. Ingemann, ed. *Proceedings of the 1982 Mid-America Conference*, Lawrence, Kansas: University of Kansas.
- Masterson, Deborah (1990) 'An LFG Solution of the MT Problem of Category Switching,' Paper Presented at the 7th International Conference on Korean Linguistics, Osaka, Japan.
- O'Grady, William (1990) 'The Korean Multiple Complementation Construction - A Categorical Approach,' Unpublished Paper, Calgary, Canada.

da: University of Calgary.

Pentheroudakis, Joseph (1990) 'You Can Get There from Here: Design and Implementation Issues in Machine Translation Systems,' Paper Presented at 1990 Annual Meeting of the American Nuclear Society, Nashville, Tennessee.

Pesetsky, David (1979) 'Russian Morphology and Lexical Theory,' ms. Cambridge, Mass.: MIT.

Pesetsky, David (1985) 'Morphology and Logical Form,' *LI* 16, 193-246.

Selkirk, Elizabeth (1982) *The Syntax of Words*, Cambridge, Mass.: MIT Press.

Shibatani, M. and T. Kageyama (1988) 'Postsyntactic Compounds in Japanese,' *Language* 64, 451-484.

Song, Seok-Choong (1988) *Explorations in Korean Syntax and Semantics*, Berkeley, CA: Institute of East Asian Studies, UC Berkeley.

Spencer, Andrew (1988) 'Bracketing Paradoxes,' *Language* 64, 663-682.

Williams, E. (1981) 'Argument Structure and Morphology,'

Department of Linguistics
University of Hawaii
1890 East-West Road
Honolulu, HI 96822
U. S. A.