An Analysis of It-Cleft Constructions¹

Jungjoo Yoon (Seoul National University)

Yoon, Jungjoo. 2002. An analysis of *It-Cleft Constructions*. *SNU Working Papers in English Language and Linguistics 1*, 142-160. The main purpose of this paper is to investigate *it-cleft* constructions in English and propose a non-derivational, lexicalist analysis of the constructions within the framework of HPSG. Based on Pollard & Sag's (1994) classification of strong and weak UDCs, I argue that there are two types of cleft clauses in *it-cleft* constructions, i.e., *wh-cleft-clause* and *bare-cleft-clause*, which have properties of strong UDCs and weak UDCs, respectively. In this paper, the two types of *it-cleft* constructions are accounted for by positing two different lexical entries of *be*_{cleft}. *It-cleft* constructions containing *wh-cleft-clause*, which have not been dealt with in the previous HPSG analyses, are analyzed by introducing a new feature CLE to distinguish *wh*-cleft clauses from other types of clauses. I propose that in the lexical entry of *be*_{cleft} for *wh-cleft-clause*, the CLE value has the same index as the focused phrase whereas in the lexical entry of *be*_{cleft} for *bare-cleft-clause*, the focused element and the SLASH value have the same index. (Seoul National University)

Keywords: HPSG, it-cleft constructions, CLE, strong and weak UDC, copula *be*, focus, cleft clauses.

1. Introduction

There are it-clefts in English as shown in examples below:

- (1) It was my little son [who/that heard an explosion last night].
- (2) It is Mary [who(m)/that my little son loves].
- (3) It was an explosion [my little son heard last night].

As shown in (1)-(3), it-clefts consist of *It*, the constituent immediately to the right of the verb *be* (referred to as a *focus phrase*) and the clause including *wh*-word or *that*, or *that*-less clause follows it (referred to as

¹ This paper was supported by the Grant for the Reform of University Education under the BK21 Project of SNU.

a cleft clause). One of major properties observed in this construction is that an element is extracted out of the cleft clause into the focus position. The extraction leaves a gap behind in the cleft clause. There is a categorical dependency between the gap and the focused element. Pollard and Sag (1994) classify it-clefts into two UDCs (unbounded dependency constructions): the examples including wh-word like (1) and (2) belongs to the class of the filler-gap constructions, or strong UDCs whereas that-less clause like (3) belongs to the weak UDCs. Chai and Lee (2002) propose a lexicalist approach to it-cleft constructions that contain that- or that-less cleft clauses. They insist that the cleft that-clause is not an adjunct modifying the preceding phrase but a complement. The verb be in it-cleft (hereafter be_{cleft}) is different from the copula be in that be_{cleft} selects for the expletive pronoun it as the subject, and two complements (that is, focus and cleft clause). They also suggest a new complementizer that cleft for the cleft construction, which is not a relative pronoun(relativizer) and which is different from the existing complementizer that.

Although Chai and Lee (2002)'s analyses of *it*-cleft constructions are a good trial to be dealt with in terms of HPSG analysis of the UDC, they have some problems. First, they don't deal with the cleft clause beginning with a *wh*-phrase. Second, a new complementizer that_{cleft} needs not be supposed in the analysis for the *it*-cleft construction. Third, the category that can occur in the focused position should not be restricted to NP, ADVP or PP. Therefore, the lexical entry of *be*_{cleft} they suggest should be modified.

In this paper, I deal with cleft clauses beginning with a *wh*-phrase as well as *that*- or *that*-less cleft clauses. I propose that the cleft clause is classified into two types, *wh-cleft-clause* and *bare-cleft-clause*, and that they can be analyzed by positing two lexical entries of be_{cleft.} Further, it is claimed that *wh*-word like *who(m)* or *whose* in *wh-cleft-clause* has non-empty CLE value. Unlike the REL, QUE, or SLASH value, the CLE value is assumed not to be amalgamated by a lexical head and obeys the CLE Inheritance Constraint.

This paper is organized as follows. In the second section, I provide a brief overview of current HPSG analyses of extraction phenomena and some related constraints, which are necessary for a proper understanding of *it*-cleft constructions. Then I briefly sketch Sag's (1997) classification of phrasal types that are cross-classified with clause types in the type hierarchy. In section 3, I show *it-clefts* analyses by Pollard and Sag (1994)

and Chai and Lee(2002), and then I review them critically. In section 4, I show my analysis in which *wh-cleft-clause* and *bare-cleft-clause* are discussed in the constraint-based theory. Finally, section 5 concludes the paper.

2. Background

2.1 Extractions

UDCs are the constructions where one element is extracted from an original place into other place, and where there is a syntactic dependency between the extracted element and the gap position. The UDCs are classified into two types: strong UDCs and weak UDCs. In strong UDCs, there is an overt constituent in a nonargument position - either a topic or an expression containing a *wh*-phrase that can be thought of as strongly associated with the gap or trace. In weak UDCs, there is no overt filler and relation between the argument and the trace is treated as one of coindexing, HPSG analysis of UDCs does not suppose an empty category such as trace or empty operator. SLASH-based analysis would suffice to explain the UDCs. All the local information of the gap position is encoded in the lexical information of the lexical head.

The UDCs have three parts: the bottom where the dependency is introduced, the middle where the information that there is an unbound trace is propagated up the tree, and the top where the SLASH value is bound off or discharged.

A synsem element of type *gap* is subject to the constraint that identifies its LOCAL value with the single member of its SLASH set:

$$(4) \quad gap\text{-synsem} \Rightarrow \begin{bmatrix} \text{LOCAL} & \boxed{1} \\ \text{SLASH} & \boxed{1} \end{bmatrix}$$

In the bottom of the extraction construction, the lexical head introduces the information on the missing element, containing an argument of type *gap-ss* in the ARG-ST value, not in the COMPS value. This is achieved in terms of the Argument Realization Principle.

(5) Argument Realization Principle (ARP)

$$word \Rightarrow \begin{bmatrix} SS \mid LOC \mid CAT \end{bmatrix} \begin{bmatrix} SUBJ \quad A \\ SPR \quad B \\ COMPS \quad C \quad elist \quad (qap-ss) \\ ARG-ST \quad A \oplus B \oplus C \end{bmatrix}$$

This principle says that the ARG-ST value is basically earned from appending the SUBJ, SPR, and COMPS list. But the *gap-ss* complement cannot be included in the COMPS value, while it can be included in the ARG-ST value.

HPSG analyses of extraction involve feature specifications for the feature SLASH that are projected upward in a syntactic structure. Extraction is thus treated entirely in terms of the inheritance of SLASH specifications, with binding off of the SLASH specification occurring at an appropriate point higher in the structure.

Following Sag (1997), I assume that words are subject to a constraint that defines their SLASH values of the members of their ARGUMENT-STRUCTURE (ARG-ST) list. English words amalgamate the SLASH values of all their arguments, obeying the constraint that is stated as in (6).

(6) Lexical Amalgamation of SLASH:

$$word \Rightarrow \left[egin{array}{lllll} ARG-ST & <[SLASH & 1 & \cdots & SLASH & n & \\ SLASH & (& 1 \cup & \cup & n & \\ BIND & & 0 & & \end{array} \right]$$

If verb's complement is slashed, the verb itself is slashed. The statement of the inheritance of SLASH specifications can be simplified in terms of the following constraint on *head-nexus phrases*.

(7) SLASH Inheritance Principle

$$hd$$
-nexus-ph \Rightarrow $\begin{bmatrix} SLASH / 1 \\ HD$ -DTR $\begin{bmatrix} SLASH / 1 \end{bmatrix}$

SLASH Inheritance Principle is a defeasible constraint that is obeyed

by all the types of head-nexus phrase considered thus far. It guarantees that the SLASH value of a phrase is the SLASH value of its head-daughter. Note that in this analysis, a subject daughter, complement daughter, etc. never passes its SLASH value to its mother. Any SLASH inheritance that appears to be an instance of such passing, is in fact mediated by the head daughter, whose SLASH value contains that of the relevant non-head daughter.

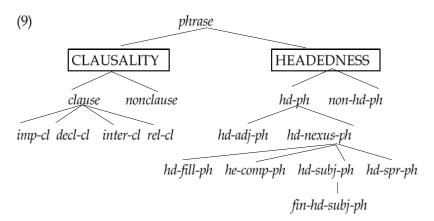
Termination of the SLASH value passed upward from the bottom is indebted to a phrasal type constraint which instances of type *hd-filler-ph* are subject to:

(8) [SLASH
$$\boxed{1} \rightarrow \boxed{LOC} \boxed{5}$$
 H phrase HEAD verb SLASH $\boxed{5} \cup \boxed{1}$

2.2 Sag's (1997) analysis

In this subsection, I briefly sketch Sag's (1997) classification of phrasal types that are cross-classified with clause types in the type hierarchy. Individual phrase types inherit both from a CLAUSALITY type and a type of HEADEDNESS. I will apply the multiple inheritance hierarchy to *it*-cleft constructions.

The mutidementional organization of phrasal types is illustrated in (9):



3. Previous analyses

3.1 Pollard and Sag (1994)

In Pollard and Sag (1994), the example of (10) belongs to the class of the *filler-gap constructions*, or *strong UDCs* whereas that of (11) belongs to the *weak UDCs*. That is, *it-clefts* are classified into two UDCs.

- (10) It's Kim who Sandy loves _____.(11) It's Kim Sandy loves _____.
- Pollard and Sag (1994) provide the examples like (12) for Binding interpretation.
- (12) a. It was herself that Mary liked best.b. It was each other that the twins liked best.

They suggest two possible analyses for anaphor binding of *it*-clefts. On one possible analysis of *it*-clefts, the relationship between the focus and the gap is analogous to that between a filler and a gap in a strong UDC. On this analysis, the SUBCAT list for the examples in (12) is as shown in (13):

On the alternative analysis, the relationship between the focus and the gap is analogous to a weak UDC in the sense that the SUBCAT list for be (in *it*-cleft) would not identify the LOCAL values of the focus and the INHER | ASH value, but instead would merely require them to be coindexed, as shown in (14):

Pollard and Sag (1994) suggest that two SUBCAT lists of *be* stem from minor idiolectal differences in the lexical entry for the copula *be* involved in the cleft construction,

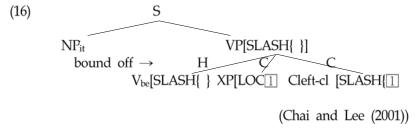
3.2 Chai and Lee (2001)

Chai and Lee (2001) show a lexicalist analysis of the it-cleft constructions in the framework of HPSG. They provide the evidence that the cleft *that*-clause is not an adjunct modifying the preceding phrase but a complement. The verb be_{cleft} selects for the expletive pronoun it as the subject, and two complements. That is, they capture the properties of the it-cleft constructions in terms of the verb be_{cleft} , of which the lexical entry is represented in (15).

(15) Lexical entry of be_{cleft}

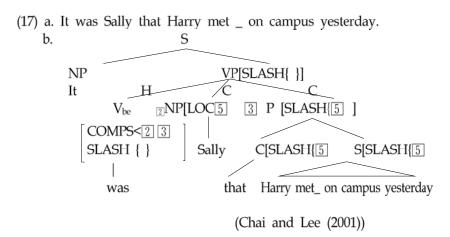
The LOC(AL) value of the first complement is structure-shared with one of the SLASH values of the second complement. Chai and Lee note that the verb *be* does not amalgamate the very SLASH value, {[3]}, from its second complement, in terms of lexical idiosyncrasy. They also note that the HEAD value of the second complement is of type verbal, subtype of comp(lementizer) and verb. This ensures the second complement could be a CP or an S.

Given such information, the *it*-cleft construction is given a syntactic structure below:



In this structure, the cleft clause is the second complement of the verb be, and a sister to the other complement, XP. The category of XP could be noun, adverb, or preposition, and the second complement is missing at least one element and its SLASH value is structure-shared with the LOC value of the preceding complement. The cleft clause, the second complement, may or may not contain the complementizer that, because it could be a CP or an S according to the lexical entry of the verb be_{cleft}.

Consider the sentence in (17a), where a non-subject noun phrase is in the focused position. The structure is illustrated as in (17b).



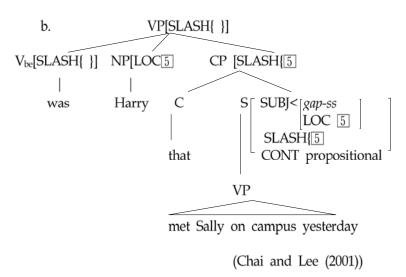
The information on the slashed argument is introduced in the ARG-ST of the lowest head verb *met* in terms of the ARP. The information is passed upward to the S, C, and CP nodes continuously by the interaction of the SLASH Inheritance Principle and SLASH Amalgamation Constraint. Then the SLASH value is bound off at the VP node, since the higher head verb was does not amalgamate the SLASH value due to its lexical idiosyncrasy.

They note that the second complement CP is of type *cp-cl*, a subtype of clause and *hd-comp-ph*, which is already exploited to account for complementizer clauses by Ginzburg & Sag (2000:48). Instances of type *cp-cl*, are subject to the constraint in (18).

(18)
$$cp$$
- cl : $[] \rightarrow H[HEAD comp], \cdots$

This constraint says that the head daughter of clauses of type *cp-d* should be a complementizer, i.e. *that* , *for*, or *to*. Since that_{cleft} is also a complementizer, *it*-clefts are subject to this constraint. The sentences, where the subject NP is extracted to be focused, are also analyzed in the same way. The verb phrase of the sentence in (19) is assigned the structure in (19b).

(19) a. It was Harry that _ met Sally on campus yesterday.



In this structure, the lowest head verb *met* contains the information that its subject is slashed, and passes it upward to its mother node, S. The information of S is exactly same with that of the VP except that it is a proposition-denoting clause while the VP daughter is a soa-denoting phrase. Note that the S is of type *decl-ns-cl*, a subtype of *decl-cl* and *hd-only-ph*, which is already used in the analyses of subject extraction by Bouma et al (2001) and of the *to-*phrase with the unexpressed subject by Ginzburg & Sag (2000). Instances of type *decl-ns-cl* are subject to the constraint in (20)

(20) decl-ns-cl: [SUBJ < [] >] \rightarrow ...

3.3 Critical Reviews of Chai and Lee (2001)

In this subsection, I review Chai and Lee's analysis critically. First of all, they limit *it*-clefts into the form of *it is/was* ~ ~ They point out that *that* can alternate with *who* or *whom* when the focused element is a human noun, and sometimes may be deleted in the cleft clause. But their analysis cannot be applied to the form of *it is/was* ~ ~ although *that*-less (*it*-cleft) construction can be covered within their analysis.

It is well known that in the form of *it is/was* ~ ~, the focus and the gap are members of the same category. However, the case that the Comp of the cleft clause contains a *wh*-word (only *who*, *whom*, and *whose* are available) is excluded.² In that case the focus and the *wh*-phrase have the same index.

- (21)a. It was Smiley_i who_i spilled beer on this couch.
 - b. It must have been Dorothyi who(m)i Fred was referring to.
 - c. It is Martha_i whose_i work critics will praise.

In the prepositional phrases containing these noun phrases, the focus and the pied-piped *wh*-phrase also have the same index.

- (22) a. It was Smiley_i on whom_i the sheriff placed the blame.
 - b. It is Margareti on whosei shoulders the burden will rest.

Chai and Lee (2001)'s analysis where *be* selects two complements, i.e. XP and one clause including a slashed XP, is acceptable in that respect. However, they don't deal with some phenomena, such as coindexing of the focus NP and a *wh*-word, or an extraction out of the cleft clause beginning with *who(m)* or *whose*.

Second, in the lexical entry of be_{cleft} in (15), they suggested "HEAD $n \vee adv \vee p$ " in the first complement. This means that the category that

² Other wh-words that serve well in restrictive relatives are not as natural in it-clefts:

i. a. ?It is this car [which I want you to sell]

b. ?It was on Thursday [when the schedule was announced].

c. ?It was in Boston [where they held the tea party]

can occur in the focused position is restricted to NP, ADVP³, or PP. But it is not true. The following examples prove that quantifier, particle, adjective or clause can occur in the focused position.

Quantifier

- (23)a. It is how much farther we have to travel that I want to know.
 - b. How far is it that we have to go?
 - c. How long was it that we stayed under water?
 - d. It is precious little that we can expect for our tax money these days.

Particle

- (24) a. It wasn't on that he pulled his boats, it was off.
 - b. It isn't in that one takes garbage, it is out.

Adjective4

- (25)a. It was happy that he looked when I saw him last.
 - b. It is happy that we all most want to be.
 - c. It was green that he painted his boat.
 - d. It was red that we sprayed the sportscar.

Clause (=CP)

- (26) a. It is that Bill will ever be this late again that is unlikely.
 - b. It is that Fred took Mary to the movies that bothers me.
 - c. It was that he had stumbled onto an anthill that Fred regretted.

(taken from Delahunty (1984))

3 Individual adverbs such as *carefully* and *regretfully* are not acceptable in the focus position, even though prepositional phrases that express the same ideas are acceptable: (C.L. Baker (1989))

a. *It was *carefully* that Donna removed the wrapping _____.

(Compare: It was *with care* that Donna removed the wrapping _____.)

b. *It was *regretfully* that Joe fired Pete.

(Compare: It was *with regret* that Joe fired Pete.)

4 Some adjective phrases cannot appear in the focused position as shown in (i).

i. *It is fond of Martha that Harry seems to be __.

Rather, the category that cannot appear in the focused position is VP as shown in (27).

(27) *It was to see his brother that Harry tired __.

Third, Chai and Lee (2000) propose to assume a new complementizer that_{CLEFI}, which selects for a finite clause which is missing a constituent. Therefore, a finite verb projection, ie. S or VP, has a nonempty SLASH value. But the property that a finite clause has a nonempty SLASH value isn't found only in cleft clauses. This property can be also found in relative clauses. Therefore, this property cannot be a motivation for a new complementizer.

4. My analysis

Based on Pollard & Sag's (1994) classification of strong and weak UDCs, I propose that there are two types of cleft clauses in *it-*cleft constructions, i.e., *wh-cleft-clause* and *bare-cleft-clause*. Each type is exemplified in the bracketed part of the following examples:

(28)	It's	Kim	[who/that5/whose mother Sandy likes].
(29)	It's	Kim	[Sandy likes].	

In (28), the *wh-cleft-clause* exhibits the characteristics of strong UDCs, while in (29) the cleft construction has properties of weak UDCs.

In order to account for the two types of *it*-cleft constructions, I will propose two different lexical entries of *be in cleft constructions*.

First, consider the example (28), which contains a wh-cleft-cl. The

⁵ Following Sag (1997), I assume *that* to be a kind of cleft *wh*-phrase. According to Sag, *that* in a relative clause can be treated as a *wh*-relative word since it is similar in most respects to relative *who* (though the latter differs in that it is in general restricted to animate antecedents).

 $[\]label{eq:continuity} i \quad \text{It is Kim} \left[\begin{array}{c} \text{on whom} \\ \text{*on that} \\ \text{*on who} \end{array} \right] \quad \text{Sandy relies.}$

lexical entry of be for the wh-cleft clause is as shown in (30).

(30) Lexical entry of becleft for wh-cleft-cl

It is assumed that wh-word like who(m), that, or whose in wh-cleft-cl has non-empty CLE value, indexed as 2 in (30). In the lexical entry of be_{cleft} for wh-cleft-clause, the CLE value have the same index with the focused element. Unlike the REL, QUE, or SLASH values that are amalgamated by HEAD, the CLE value obeys the constraint that is stated as in (31)

- (31) CLE Inheritance Constraint⁶ In a headed phrase,
- a) if the value is of sort preposition, the CLE value is inherited from the complement daughter's CLE.
- b) otherwise, the CLE value is inheritied only from a leftmost daughter's CLE.

The other constraint that wh-cleft clauses inherit is shown in (32). This constraint is the general constraint governing all clauses.

(32) clause
$$\Rightarrow$$
 SUBJ list (PRO)
HEAD [MOD/none]
REL { }
QUE { }

⁶ This constraint is similar to the constraint on Interrogative Pied piping suggested by Pollard and Yoo (1998:435) in that the given value is inherited from the leftmost element like specifier daughter..

likes

The constraint in (33) says that the CLE value of non-head-daughter must be non-empty. The CLE value of *wh-cleft-cl* is inherited from the CLE value of the leftmost non-head-daughter.

Since non-finite cleft clauses are not allowed, cleft clauses have the following constraint:

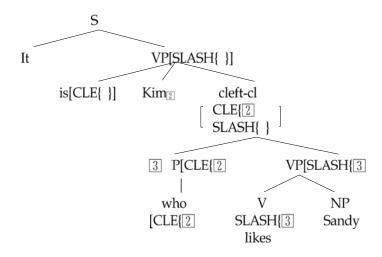
(34)
$$cleft-cl \Rightarrow VFORM fin$$
]

The constraint in (34) says that cleft clause must be a finite clause. The examples of *wh-cleft-cl*, interacting with the above constraints, are illustrated in (35) and (36).

(35) a. It is Kim whose mother Sandy likes. b. VP[SLASH{ }] It is[CLE{ }] Kim₂ cleft-cl $CLE{2}$ SLASH{} 3 P[CLE{2 VP[SLASH{3 NP V mother whose $[CLE{2}]$ SLASH[3] Sandy

156 Yoon, Jungjoo

(36) It is Kim who likes Sandy.

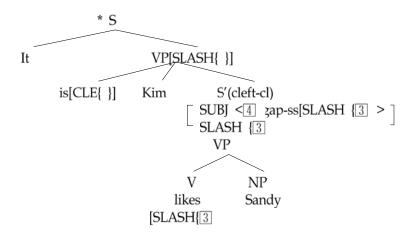


Now consider the second type of lexical entries of be_{cleft}. The lexical entry of be_{cleft} for the cleft construction with *bare-cleft-clause* is shown in (37).

(37) Lexical entry of be_{cleft} for bare-cleft-cl

This constraint says that the focused element and the SIASH value have the same index. Note that (37) requires that the subject in the second complement must not be extracted (via the [SUBJ < >] specification). Therefore, ungrammatical examples like *It is Kim likes Sandy will not be allowed.

(38) * It is Kim likes Sandy.



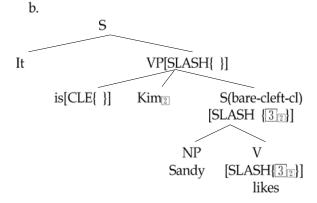
The second type of cleft clauses like (29) is treated in terms of both *bare-cleft-cl* and *fin-hd-subj-ph* and hence inherits constraints from both of these supertypes. The constraint is shown in (39):

(39) bare-cleft-cl
$$\Rightarrow$$
 $\begin{bmatrix} SUBJ < > \\ SLASH \{[\]\} \\ CLE \ \{\ \} \end{bmatrix}$

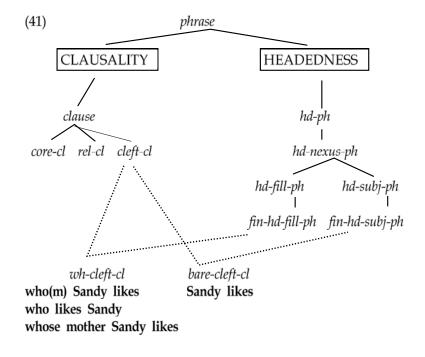
The constraint in (39) says that the subject must exist in the bare-cleft-clause as shown in SUBJ < >. The cleft clause has a non-empty SLASH value. Without wh-word, CLE value of cleft clause is empty. The bracketed part in (29), interacting with the above constraints, is sketched in (40).

158 Yoon, Jungjoo

(40) a. It is Kim Sandy likes.



To sum up, the types of cleft clauses discussed in this section are illustrated in (41).



5. Conclusion

In this paper, I propose a non-derivational, lexicalist analysis of it-cleft constructions within the framework of HPSG. Based on Pollard & Sag's (1994) classification of strong and weak UDCs, I argue that there are two types of cleft clauses in it-cleft constructions, i.e., wh-cleft-clause and bare-cleft-clause, which have properties of strong UDCs and weak UDCs, respectively. In this paper, the two types of it-cleft constructions are accounted for by positing two different lexical entries of be_{cleft}. It-cleft constructions containing wh-cleft-clause, which have not been dealt with in the previous HPSG analyses, are analyzed by introducing a new feature CLE to distinguish wh-cleft clauses from other types of clauses. In the lexical entry of be_{cleft} for wh-cleft-clause, the CLE value has the same index as the focused phrase whereas in the lexical entry of becleft for bare-cleft-clause, the focused element and the SLASH value have the same index. I propose the CLE value obeys the CLE Inheritance constraint, unlike the REL, QUE, or SLASH values that are amalgamated by HEAD.

References

Akmajian. A. 1970. On deriving cleft-sentences from pseudo-cleft sentences. LI 1. 149-68.

Baker, C. L. 1989. English Syntax. MIT Press.

Bouma, Gross, Robert Malouf, and Ivan Sag. 2001. Satisfying Constraints on Extraction and Adjunction. *Natural Language and Linguistic Theory* 19 1-65..

Chai, Myong-Hi and Lee, Namgeun. 2001 A lexicalist Approach to It-cleft Constructions. Studies in Modern Grammar 26, 67-85.

Davidse, Kristin. 2000 A constructional approach to clefts. Linguistics 38-6, 1101-1131.
Gerald P. Delahunty. 1984. The analysis of English cleft sentences. Linguistic Analysis
13. 63-113

Ginzburg, J. and Sag I. English Interrogative constructions. ch.5.

Gundel. J.K. 1977. Where do cleft sentences come from? Lg. 53. 543-59

Hankamer, Jorge. 1974. On the non-cyclic nature of WH-clefting. Papers from the 10th Regional Meeting, *Chicago Linguistic Society*. 221-32.

Kim Jong-Bok and Byung-Soo Park. 1996. English free relative clause constructions: from a constraint-based perspective. *PACLIC* 11. 31-41.

160 Yoon, Jungjoo

Pinkham, Jessie, and Jorge Hankamer. 1975. Deep and shallow clefts. Papers from the 11th Regional Meeting, *Chicago Linguistic Society*. 429-50.

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

Schachter, Paul. 1973. Focus and relativization. Lg. 49. 19-46.

Stahlke, Herbert F. W. 1976. Which that. Lg. 52. 584-610.

Sag, Ivan 1997. English relative clause constructions. Linguistics 33, 431-483.

Sag, Ivan and Thomas Wasow 1999. Syntactic Theory: A Formal Introduction. CSLI Publications.

Iungioo Yoon jungjoo@hotmail.com