Syntax and Semantics of English It-Cleft Constructions: A Constraint-Based Analysis

Jong-Bok Kim
(KyungHee University)

Jong-Bok Kim. 2007. Syntax and Semantics of English It-Cleft Constructions: A Constraint-Based Analysis. ***Studies in Modern Grammar*** **48, 217-235.** The so-called it-cleft construction in English packages information structure in a special way. In particular, it places 'salient' discourse information in the postcopular position. This paper examines the syntactic, semantic, and pragmatic properties of the construction, and provides a constraint-based analysis of these properties within the framework of HPSG.

Key words: it-cleft constructions, discourse information, expletives, extraposition, focus

1 Basic Properties

1.1 Distributional Properties

The English it-cleft sentence, exemplified in (1), is a complex sentence consisting of a matrix clause headed by the copula *be* and a relative-like cleft clause:

(1) It is the syntax teacher that we are looking for.

As noted by Lambrecht (2001) and others, the it-cleft embodies a discourse function by placing a focused (or pivot) element in the postcouplar position and a presupposition in the cleft clause. To be more precise, (1) is taken to have the following information structure:

(2) a. Presupposition: We are looking for x.

b. Foreground: the syntax teacher

c. Assertion: x is the syntax teacher

As for the possible phrasal type for the foreground or the pivot phrase, we observe that only a limited type of phrases is possible:[1]

(3) a. It was [the gauge] that was the killer in the first place . (ICE-GB)

b. AdvP: And it was [then] that he felt a sharp pain. (ICE-GB)

c. PP: It was [in Berilin] that he first read Morelli's work. (BNC)

d. Subordinate Clause: It wasn't [till I was perhaps twenty-five or thirty] that I read them and enjoyed them (ICE-GB)

Canonically, verbal phrases like AP or VP cannot occur as the pivot phrase as in (4) though literature provides some examples like (5) (cf. Declerck 1984, Huddleston and Pullum 2002):

(4) a. *It is fond of John that Harry seems to be.

b. *It was to see his brother that Harry tried.

(5) a. It's [certainly not to make life easier for us] that they are changing the rules.

b. It wasn't [green] I told you to paint it.

There also exist it-cleft examples with no pivot XP (from the BNC).

(6) a. It must be that you also perceive it.

b. It could be that your lack of confidence is connected with work.

c. It might be that they could be only reached through their elders.

When the highlighted XP is a clause, only a subordinate clause is possible, but no content clause can occur as the pivot XP:[2]

(7) a. It is because you stood up for yourself that you were sacked.

b. It was what you put it and what you achieve which counts. (ICE-GB)

[1]The bracket is adopted for the readability and examples drawn from the International Corpus of English are marked as ICE-GB whereas those from the British National Corpus as BNC.

[2]As noted in Prince (1978), it is however, well-formed when the content clause functions as the subject as in *It is that Kim snores that bothers me.*

(8) a. *It's that he did it deliberately that I'm inclined to think.

b. *It is why no one told us that I'm wondering.

What we can observe here is that except for verbal elements and content clauses, various types of phrases including subject, complement, and adjunct can function as the pivot phrase.

1.2 Semantic and Pragmatic Properties

As noted by literature (cf. Gundel 1977, Prince 1978, Collins 1991), one salient property of the cleft constructions is that it represents the background information. This can be attested by its preservation under the negation.

(9) a. It was the teaching material that we used.

b. It wasn't the teaching material that we used.

Both the positive and negative sentences convey the information that we used something x. This variable x is what the pivot phrase *the teaching material* expresses. That is, while the background information conveys an open proposition with a variable x, and the foreground provides the value of this variable. This value is in general exhaustive and exclusive, inducing a contrastive meaning:

(10) In fact it's their teaching material that we're using (ICE-GB)

Here, the sentence means we are not using his or her teaching material, but their teaching material.

As noted by H & P (2002), the presupposition expressed by the background part can be either discourse-old or discourse-new. Observe the following direct conversation extracted from the ICE-GB:

(11) A: And hit it as far as you can go on [the top note] but still think and think very hard on your palate
B:...
A: It's a case of keeping a bit more open sort of the jaw more unhinged and the palate up like crazy the whole time and just simply as if you loosen all the nuts and bolts on the valve
B: Yes
A: Watch coming down,..
B: ...

> A: It's [that note] that's a little flat because you've stopped concentrating on making these higher notes and you do let the whole thing sag a little

As shown in the dialogue here, the highlighted XP *that note* is not discourse new. It is the cleft-clause that introduces discourse new information as observed from the last dialogue. As noted by Prince (1978) and others, in it-clefts, foreground thus does not necessarily carry focus: when the background is discourse-new, it is usually the focus that is discourse old. This means that there are at least two different types of it-clefts, depending on the position of the focus.[3]

2 Previous Analyses

2.1 Extraposition Analyses

Among several main approaches for the it-cleft construction, the extraposition analysis assumes a direct syntactic or semantic relation between the cleft pronoun *it* and the cleft clause through extraposition (Akmajian 1970, Emonds 1976, Gundel 1977, Hedberg 2000). For example, Gundel (1977) assumes that the it-cleft (12)c is derived from (12)b which is derived from the pseudo-cleft (12)a:

(12) a. What you heard was an explosion. (pseudocleft)

b. It was an explosion, what you heard. (right-dislocated pseudocleft)

c. It was an explosion that you heard.

The structure that Gundel adopted can be represented as the one in (13):

[3]Collins' (1991), examining the London-Lund and the Lancaster-Oslo/Bergen corpus, claims that the it-cleft is 'newness-oriented' whereas the wh-cleft is 'giveness-oriented', with the inverted wh-cleft sharing both features.

In addition, only the cleft clause of it-clefts can have the PP wh-head:

(17) a. And it was this matter [on which I consulted with the chairman of the Select Committee].

b. *[On which I consulted with the chairman of the Select Committee] was this matter.

The lack of such isomorphic relations among the two clefts indicates that we cannot derive one from the other or vice versa.

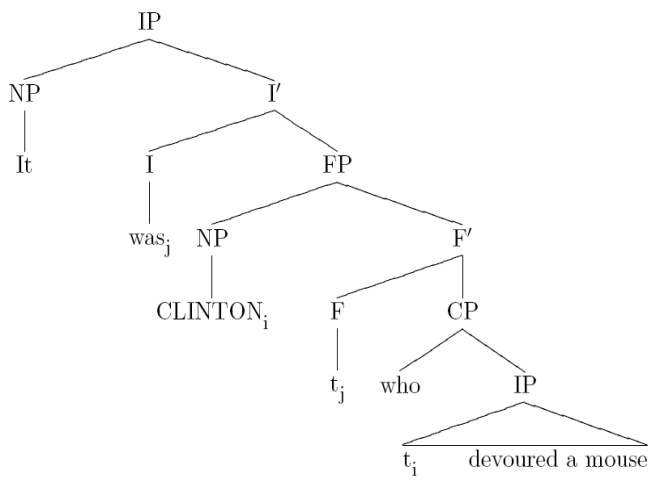
2.2 Expletive Anlayses

Meanwhile, the expletive analysis (Heggie 1988, Kiss 1998, Chomsky 1977, Huddleston 1984, Delahunty 1982, Delin 1989, É, Kiss 1998, Lambrecht 2001) takes the pronoun *it* to be an expletive expression with no semantic meaning while the cleft clause is semantically linked to the clefted constituent by a predication relation.

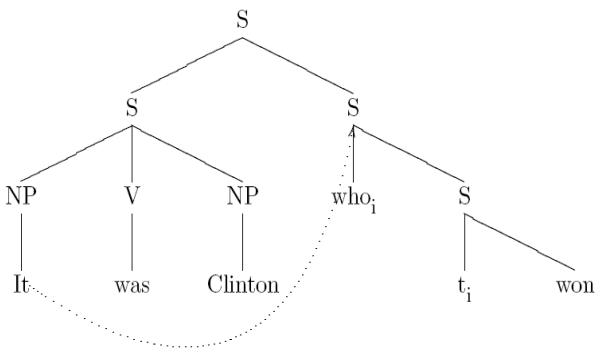
(18) It was [John + that I saw].

An elaborated analysis within this view has been proposed by Kiss (1998). The cleft constituent is moved from an embedded IP to the Spec of the FP:

(19)



(13)



As represented here, the it-clefts are reduced forms of right-dislocated pseudo clefts, in which *it* functions as a pronominal reference to the topic which appears at the end of the sentence.

Even though the two are identical in presenting 'salient' discourse information in a special way, they have different syntactic properties which make it hard to drive one from the other (cf. Pavey 2004). For example, one noticeable difference lies in that only the wh-cleft allows bare infinitives as the highlighted XP phrase:

(14) a. What you should do is [$_{\text{VP}}$ order one first].

b. *It was [$_{\text{VP}}$ order one first] that you should do first.

In addition, unlike the it-cleft, the wh-cleft does not allow an adverbial subordinate clause as its pivot clause:

(15) a. It wasn't [till I was perhaps twenty-five or thirty] that I read them and enjoyed them.

b. *When I read them and enjoyed them was [not until I was perhaps twenty-five].

The possible types of cleft clauses are also different. For example, unlike it-clefts, neither wh-clefts nor inverted wh-clefts allow the cleft clause headed by *that*:

(16) a. It's the writer [that gets you so involved].

b. *[That gets you so involved] is the writer.

As represented here, Kiss (1998) claims that the clefted phrase, functioning as identificational focus, occupies the specifier of the FP through a movement process whereas *it* and the copular verb *be* are introduced as expletives. Within traditional wisdom, it is hard to assume that pronouns like *that* and *this* have no semantic meaning at all.

As noted by Hedberg (2000) and Davidse (2000), the problem for such an analysis can have may arise from constructions like the following:

(20) a. There's Jim who makes the coffee.

b. This is not Iowa we're talking about.

c. That's the French flag you see flying over there.

As such examples, the cleft noun *it* can be replaced by *this* or *there* or *that*, depending on pragmatic constraints (cf. Hedberg 2000).

Another possible issue that can arise in such an anlaysis is that the pivot phrase does not always carry focus value. Even there are examples that only part of the pivot phrase is focused:

(21) It was Jack and TOM that visited Mongolia.

It is not difficult to construct a context in which of the NP *Jack and TOM*, only TOM gets focus but Jack is presupposed.

3 A Constraint-Based Analysis

3.1 Two types of it-cleft constructions

As noted by Heycock and Kroch (1999), the copula in the cleft construction is 'specificational', not 'predicational'. In sentences like *John is happy*, the copula is used as predicational, whereas in sentences like *The culprit is John*, the copula is specificational. One main difference is that in the former the postcopular element denotes the property of the subject whereas in the latter it denotes an individual. In terms of the argument structure, this indicates that the cleft copula *be* selects two arguments which refer to the identical individual:

(22) $\begin{bmatrix} \langle \text{be} \rangle \\ \text{ARG-ST} \left\langle \text{NP}_i, \text{XP}_i \right\rangle \end{bmatrix}$

These two arguments of the specificational *be* will canonically be realized as SPR (specifier) and COMPS (complements) in syntax:[4]

(23) Canonical Argument Realization of *be*:

$$\left[\text{ARG-ST}\left\langle \boxed{1}\text{NP}_i,\ \boxed{2}\text{XP}_i \right\rangle\right] \quad \Rightarrow \quad \begin{bmatrix}\text{SPR}\left\langle \boxed{1}\text{NP}_i \right\rangle \\ \text{COMPS}\left\langle \boxed{2}\text{XP}_i \right\rangle\end{bmatrix}$$

Such an argument realization will generate canonical specificational sentences like the following:

(24) a. The recipient of this year's award is President Kim.

b. The one who broke the window was Mr. Kim.

However, in cleft-constructions, the arguments are realized in different ways, depending on how the information structure (IS) is realized. That is, the three types of clefts in (25) reflect how the arguments are realized differently with respect to the information structure of the sentence in question:

(25) a. it-cleft: It is that syntax book that we are looking for.

b. wh-cleft: What we are looking for is that syntax book.

c. inverted-cleft: That syntax book is what we are looking for.

As we have seen, in the it-cleft the pivot XP is canonically focus, whereas in the wh-cleft it is the postcopular phrase that functions as focus and in inverted-cleft, the final cleft clause is focus.

Two common information structure sensitive features are TOPIC and FOCUS, which are usually linked to given and new information, respectively. In addition to these two features, we introduce the feature HIGHLIGHT. The feature HIGHLIGHT is similar to the notion of 'salient': the information that is most salient in the given context bears this feature. Consider the following simple question and answer dialogue:

(26) A: What did John drink?
B: John drank beer.

[4]The boxed integers are used to represent the structure sharing between two values.

It is clear that in the expressions 'John' and 'drank' here are both given information (topic), whereas 'beer' is new information (focus). The difference between 'John' and 'drank' is just that 'John' is more salient than 'drank' since it is what the sentence is about. This kind of comparison also holds between completive (pure) focus and contrastive focus:

(27) A: Did John drink beer or coke?
B: John drank beer.

Unlike the NP 'beer' in (26), the NP 'beer' here is focus, but has a contrastive meaning compared to 'coke'. In this sense, we take 'beer' to be contrastive focus, the most salient information in this given discourse. The feature HIGHLIGHT is thus given to the topic and contrastive focus. The feature thus can be assigned either to a TOPIC or to a FOCUS expression.

This kind of system that introduces the feature HIGHLIGHT to mark the discourse salient information allows us to assign the FOCUS value even to the cleft clause as we have seen in (11). Examples like (28) illustrate the same point (Huddleston and Pullum 2002):

(28) a. It was [50 years ago] that the first real computer was invented.

b. It is [with great pleasure] that I now declare this Exhibition open.

In such examples, it is not the expressions *50 years ago* and *with great pleasure* but the cleft clause that conveys new information. In the current system, the adverbial elements get the value of the feature HIGHLIGHT whereas the cleft clause FOCUS.

Also notice that there are two different types of it-clefts. Compare the following:[5]

(29) a. Type A: It is [$_{\text{PP}}$ on Bill] [$_{\text{CP/PP}}$ that John relies __].

b. Type B: It is [$_{\text{NP}}$ Bill] [$_{\text{S}}$ [on whom] [John relies]].

In (29a), the cleft clause contains a gap matching with the filler PP *on Bill*. Even if we treat *that* as the relativizer, the PP gap cannot be discharged by *that*. However, in (29b) the cleft clause has two parts: one with a missing gap *John relies* and the other with the wh-phrase *on whom* functioning as the filler. These two make the cleft clause a complete sentence. This second type is similar to examples where the highlighted element is an adverbial element:

[5]See Gazdar et al. (1985) and Kim and Sells (2007) recognizing two different *it*-cleft constructions.

(30) a. It was [then] when we all went to bed.

b. It was [only gradually] that I came to realize how stupid I was.

In addition, as observed in the literature, even though the cleft clause is similar in structure to a restrictive relative clause, there are several considerable differences. For example, consider the following:

(31) a. It is John that we are looking for.

b. *John that we are looking for showed up.

Unlike the cleft, the canonical restrictive relative clause does not allow a pronoun to function as the antecedent of the relative clause.

In addition, note that in cleft clauses not all wh-phrases can freely occur even though they introduce *who*, *whose*, and even zero:

(32) a. It's the second Monday [that] we get back from Easter holiday.

b. Perhaps it was the peasant girl [who] got it

c. It's in the scenes [when] De Niro fighting against an on-rush of uncoordinated tics and twitches is beginning to relapse into the coma.

d. It is Uncle John [whose address] I lost.

To capture these two different types and restrictions on the type of wh-phrases, we first assume that both are used to highlight the contrastive focus, but different with respect to how the arguments of the specificational *be* are realized.[6]

(33) Argument Realization for Type A It-Cleft:

$$\Big[\text{ARG-ST } \langle \text{XP}_i, \boxed{1}\text{YP}\rangle\Big] \quad \Rightarrow \quad \begin{bmatrix} \langle \text{be}\rangle \\ \text{SPR } \langle \text{NP}[\textit{it}]\rangle \\ \text{COMPS } \Big\langle \boxed{2}\text{YP}_i\Big[\text{HIGHLIGHT } +\Big], \text{CP}\Big[\text{GAP } \langle \boxed{2}_i\rangle\Big]\Big\rangle \\ \text{GAP } \boxed{A} \ominus \langle \boxed{2}\rangle \end{bmatrix}$$

[6]We do not assign the feature FOC in the lexical realization since its realization can be dependent upon context, even though the highlighted phrase canonically is contrastive focus.

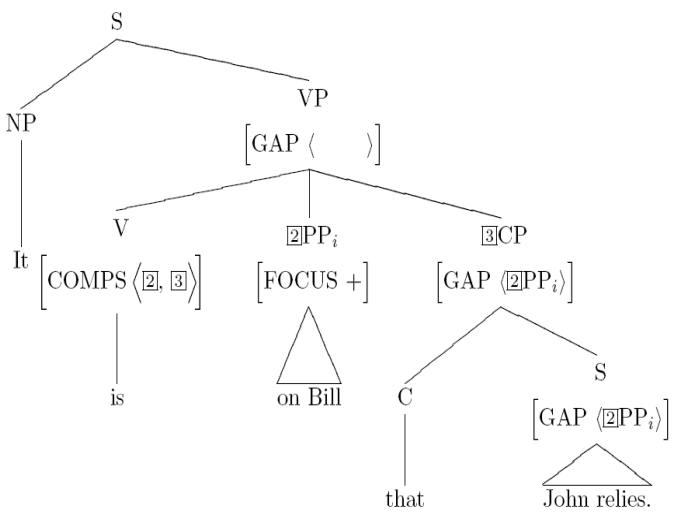
(34) Argument Realization for Type B It-Cleft:

$$\left[\text{ARG-ST } \langle \text{XP}_i, \boxed{2}\text{YP} \rangle\right] \Rightarrow \left[\begin{array}{l} \text{SPR } \langle \text{NP}[\textit{it}] \rangle \\ \text{COMPS } \left\langle \boxed{2}\text{YP}\left[\text{HIGHLIGHT } +\right], \text{S}\left[\begin{array}{l}\text{MOD } \langle \boxed{2} \rangle \\ \text{GAP } \langle \quad \rangle\end{array}\right] \right\rangle \end{array}\right]$$

In both constructions, the contrastive focus functions as the most salient contextual information. However, in Type A, the second complement, functioning as the background, is a CP with one GAP value. Notice that the index of the GAP value is identical with that of the pivot XP focus. This means that there is a strict semantic linkage between these two. Meanwhile, in Type B, the second COMPS element is a saturated S that modifies the focused element.

Let us consider the structure that (33) generates:

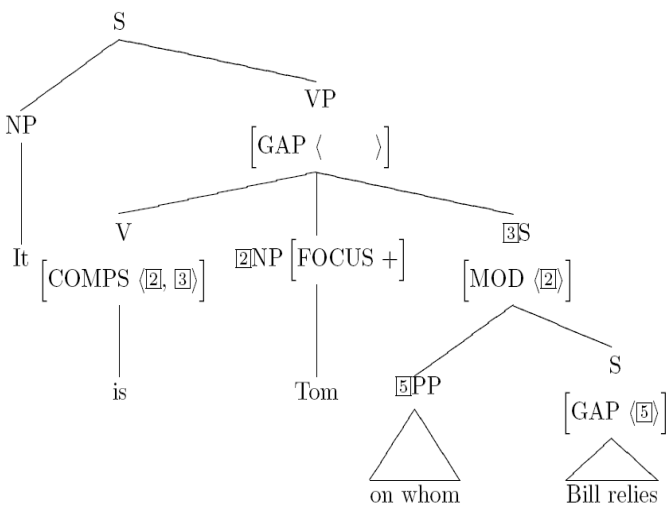
(35)



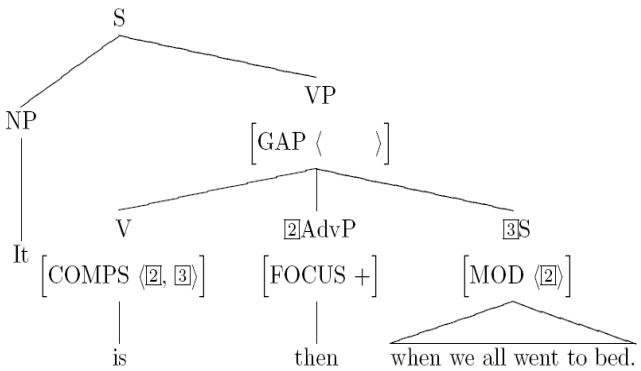
This is a Type A cleft sentence: The copula *be* selects two complements: PP and CP. The cleft clause CP has a PP gap whose GAP value which is linked to the pivot PP.

Meanwhile, the realization of (33b) will project sentences like the following:

(36)



(37)



Both sentences are Type B it-cleft constructions in the sense that the cleft clause contains no GAP element and modifies the pivot phrase.[7]

This kind of analysis will also explain why the following examples are unacceptable:

(38) a. *It is [Kim] [[on whom] [that Sandy relies]].

b. *It is [Kim] [[on whom] [Sandy relies on]].

[7]The MOD feature here is originated from the subordinator conjunction *when*.

c. *It is [Kim] [[whom] [Sandy relies]].

(38)a is ruled out since the combination of *[[on whom] [that Sandy relies]]* is not a well-formed S though it could be a CP; (38)b is not allowed because of the mismatch between the gap (NP) and the filler (PP); (38)c is ruled out in a similar sense. The sentence *Sandy relies* requires a PP but the filler is an NP (*whom*)

Within the present system where the missing element in the cleft clause is taken to be a GAP element, we can also predict an unbounded dependency relation:

(39) a. It was the director that she wants to meet __ .

b. It was the director that she said she wants to meet __ .

c. It was the director that he thinks she said she wants to meet __ .

As noted here, the distance between the pivot phrase *the director* and the putative gap in the cleft clause can be long distance.

In addition, the system allows us to generate examples like the following:

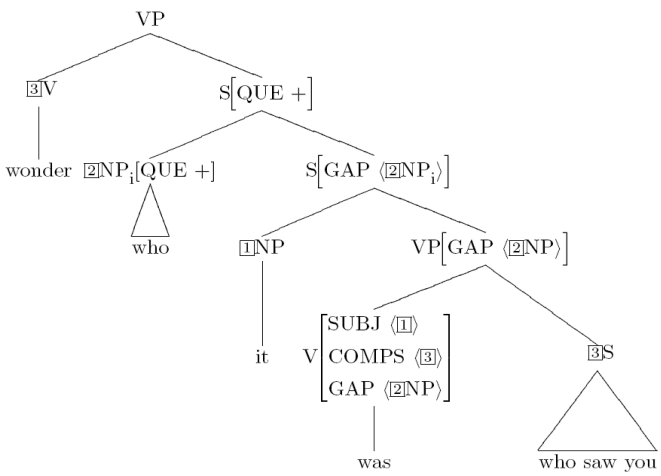
(40) a. I wonder who it was o̊who saw you.

b. I wonder who it was __ you saw __ .

c. I wonder in which pocket it was __ that Kim had hidden the jewels.

Let us then consider the structure of (40)a that our system generates:

(41)



As noted here, the first COMPS value of the cleft copula *be* is realized as a GAP element. This GAP value is passed up to the point it is discharged by the wh-element *who* which induces an interrogative meaning to the complement clause of the verb *wonder*.

Notice that even though the present system allows the pivot phrase to be gapped, the GAP value originated in the cleft clause cannot pass up further:

(42) a. Who do you think it is that Mary met?

b. *Who do you think it is John that Mary met?

Notice that in (33) the GAP value originated in the cleft clause is lexically terminated at the CP level by the copula *be*, and that in (34), the cleft clause S does not contain any GAP. The two realizations thus correctly block us from generating examples like (42)b.

3.2 Additional Formal Properties

The it-cleft also shows the so-called connectedness effect. Observe the following:

(43) a. Oscar$_i$ loved him$_{*i/j}$.

b. Oscar$_{i/j}$ loved himself$_{i/*j}$.

(44) a. It was Oscar$_{i/j}$ that loved himself$_{i/*j}$ the most.

b. It was Oscar$_i$ that loved him$_{*i/j}$ the most.

(45) a. It was him$_{*i/j}$ that Oscar$_i$ loved the most.

b. It was himself$_{i/*j}$ that Oscar$_i$ loved the most.

What we can observe here is that the binding relationships in simple sentences are carried over to cleft clauses. When the focused phrase is a pronominal that corefers with an NP in the cleft clause, then the pronominal must be a reflexive. Meanwhile, if it is a non-reflexive, it must have a different referent. This connectivity is expected within the binding theory that refers to the ARG-ST as set forth by Sag et al. (2003):

(46) Binding Principles

a. An anaphoric element must be outranked by a coindexed element in the ARG-ST.

b. A pronominal element must not be outranked by a coindexed element in the ARG-ST.

Given this principle, let's consider the ARG-ST of the verb *love*.

(47) a. $\left[\text{ARG-ST}\left\langle \text{NP}_{\text{i}}[pron],\ \text{NP}_{*\text{i}/\text{j}}[pron]\right\rangle\right]$

b. $\left[\text{ARG-ST}\left\langle \text{NP}_{\text{i}}[ana],\ \text{NP}_{\text{i}/*\text{j}}[ana]\right\rangle\right]$

The object pronominal (*pron*) occurring as the second element in the ARG-ST thus cannot be coindexed with the subject whereas the object anaphor (*ana*) must be coindexed with the subject. Notice that in the cleft sentences, the ARG-ST of *love* is not different from these. This indicates that the identical binding facts are observed in clefts,

Agreement in it-clefts also is intriguing. Observe the following data (cf. Akmajian

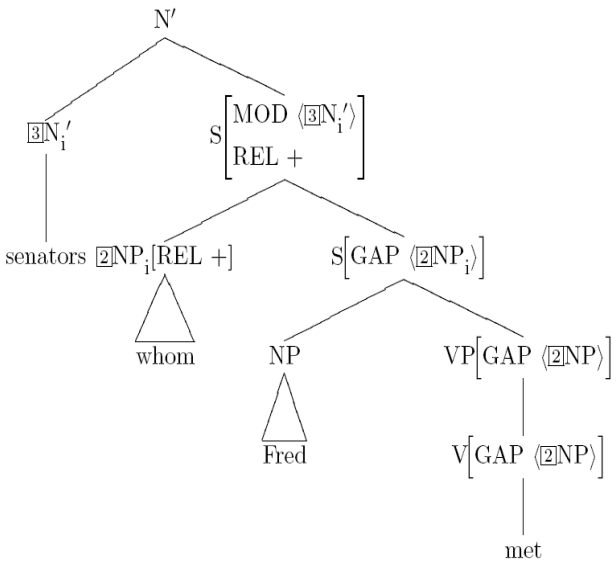
(48) a. It's me/you/him that likes/*like hand gliding.

b. It is you$_{sing}$ that likes/*like hand gliding.

c. It is you$_{pl}$ that *likes/like hand gliding.

Before we discuss these data, let us consider how to deal with canonical relative clauses within the HPSG system:

(49)

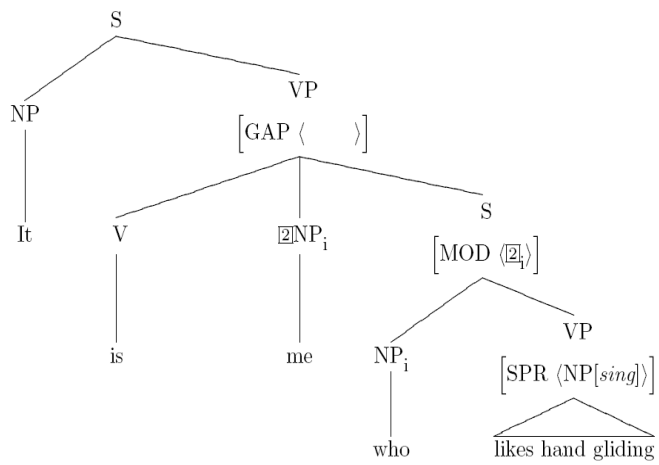


As noted here, the relative clause is modifying the head noun *senators* through the MOD feature whose index value (i) is identical with the relative pronoun *whom*. This relative pronoun is matching with the head noun.[8] Cleft-clauses are not different from relative clauses in terms of semantics. One thing to notice here is that there is thus a semantic coindexation relation between the head noun and the modifying relative clause.

This semantic coindexation relation is carried over to cleft clauses which in terms of propositional meaning are identical with the relative clauses but different only in terms of structures. Let us consider the structure for (48)a:

[8]See Sag (1997) for a detailed semantic process for relative clauses.

(50)



The VP *likes hand gliding* requires a singular subject as indicated by its SPR value. The pronoun *that* now functions as the subject of this VP. This NP's index value is identical with the index value of the pivot phrase that the cleft clause modifies. The expressions *me* and *who* share the identical number value. This system then easily accounts for the remaining examples in (49).

4 Conclusion

The English it-cleft construction marks a special information structure by a syntactic bi-clausal option. In the paper we have reviewed the basic syntactic properties of the construction as well as semantic and pragmatic properties.

Based on these observations, the paper has provided an analysis to deal with the formal properties of it-cleft constructions: constituent structures of two different types of it-clefts, connectedness effects, and agreement. We have seen that when we resort to tight interactions among various grammatical components such as argument-structure, semantics, information-structure, we can have a streamlined analysis of these properties.

References

Akmajian, Adrian. 1970. On deriving cleft sentences from pseudo-cleft

sentences. *Language* 1: 149–168.

Collins, Peter C. 1991. *Cleft and Pseudo-Cleft Constructions in English.* London and New York: Routledge.

Davidse, Kristin. 2000. A Constructional Approach to Clefts. *Linguistics* 38.6: 1101-31.

Declerck, Renaat. 1984. Some Restrictions on Clefts That Highlight Predicate Nominals. *Journal of Linguistics.* 20.1: 131-154.

Delahunty, Gerald. 1984. The Analysis of English Cleft Sentences. *Linguistic Analysis.* 13 (2): 63-113.

Delin, Judy. 1995. Syntactic Constraints on Discourse Structure: The Case of It-Clefts. *Linguistics* 33.3: 465-500.

Engdahl, Elizabet and Enric Vallduví. 1996. Information Packaging in HPSG. In Claire Grover and Enric Vallduví (eds.), *Studies in HPSG*: Edinburgh Working Papers in Cognitive Science, 1–31.

Gazdar, Gerald, Ewan Klein, Geoffrey K. Pullum, and Ivan A. Sag. 1985. Generalized Phrase Structure Grammar. Cambridge, MA; Havard University Press and Oxford; Basil Blackwell.

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

Hedberg, Nancy. 2000. The Referential Status of Clefts. *Language* 76.4: 891-920.

Heggie, Lorie. 1993. The Range of Null Operators: Evidence from Clefting. *Natural Language & Linguistic Theory.* 11 (1): 45-84.

Heycock, Caroline and Anthony Kroch. 1999. Pseudo-cleft connectedness: Implications for the LF interface level. Linguistic Inquiry 30, 365-397.

Huddleston, Rodney and Geoffrey Pullum edt. 2002. The Cambridge Grammar of the English Language. Cambridge University Press.

Kim, Jong-Bok, and Peter Sells. 2007 (to appear). *English Syntax: An Introduction.* Stanford: CSLI Publications.

Kiss, Katalin E. 1999. The English Cleft Construction as a Focus Phrase. *Boundaries of Morphology and Syntax*, 217-29.

Lambrecht, Knud. 2001. A framework for the analysis of cleft constructions. *Linguistics*, 39.3: 463–516.

Prince, Ellen. 1978. A comparison of wh-cleft and it-clefts in discourse. *Language* 54.4: 883-906.

Huddleston, Rodney, and Geoffrey Pullum. 2002. *The Cambridge Grammar of the English Language.* Cambridge: Cambridge University Press.

Sag, Ivan. 1997. English Relative Clause Constructions. *Journal of Linguistics* 33.2: 431–484.

School of English
Kyung Hee University
1 Hoegi-dong, Dongdaemun-gu
Seoul 130-701
(02) 961-0892 jongbok@khu.ac.kr

Received: March 13, 2007
Revised version: June 1, 2007