That's Something That I Wouldn't Want to Have to Account For, Is a Sentence Like This One.*

Richard Douglas Jehn

1. Introduction

The grammar of focus phenomena in English has come under intensive investigation in recent years and this research has provided much in the way of explanation for structures which had previously been little understood, e.g. cleft constructions, topicalization, etc. There is, however, at least one type of focus construction which seems to pattern like pseudo-cleft sentences, but which has been neglected in the literature. This paper outlines the most apparent aspects of the syntactic and semantic behaviour of the construction in question and presents a tentative proposal for its incorporation into the grammar of English following the framework of the revised extended standard theory (Chomsky 1975, 1977a; Chomsky and Lasnik 1977).

2. Th-Cleft and Its Similarities to Wh-Cleft

This portion of the paper presents the general syntactic behaviour of the construction in question. Some semantic distinctions which are reflected by overt syntactic markers are presented and tentative interpretations of the phenomena are offered.

In example (1) below, I present a declarative sentence as (1a), the pseudo-cleft sentence, which may (or may not) be derived from (la),1 as (1b), and the construction under investigation here, which I will term "th-cleft", as (1c):

(1) a. I won't go to the bar alone.
   b. What I won't do is go to the bar alone.
   c. That's what I won't do, is go to the bar alone.2

Although the surface structures of (1b) and (1c) are clearly related, establishing the deep structure relationship is problematic. I will return to this issue in Section 4 after presenting some general characteristics of the th-cleft construction.

A highly unusual feature of th-cleft constructions is that the wh-clause is shared over two Ss; that is, the embedded sentence in (1c) behaves as if it were both predicate NP of that is and subject of the second copula. Consider sentences (2) and (3):

(2) a. [That] is [what I never would have believed], is [that the earth is round.]
   b. [That] is [what I never would have believed] (cf. I never would have believed that.)
   c. [What I never would have believed] is [that the earth is round.] (cf. I never would have believed that the earth is round.)
(3) a. [That]'s [what I won't do], is [go to the bar alone.]
b. [That]'s [what I won't do.] (cf. I won't do that.)
c. [What I won't do] is [go to the bar alone.] (cf. I won't
go to the bar alone.)

The (b) and (c) sentences above parallel the sentences in parenthes­ses, but there are no corresponding sentences for the (a) examples. The fact that only one copula appears in the (b) and (c) sentences while two appear in the (a) sentences suggests why many speakers find them ungram­matical. Many more speakers find the construction acceptable with a pause instead of a copula preceding the focus constituent:

(4) That's what I never would have believed -- that the earth is round.

(5) That's what I won't do -- go to the bar alone.

Sentences (4) and (5) suggest that the second copula is spurious. More convincing evidence for the spuriousness of this copula is evinced by the application of question formation to the sentences. Consider sen­tences (6) and (7):

(6) Is that what he never would have believed, is that the earth
is round?
(7) *Is that's what he never would have believed that the earth
is round?

Although the main verb directly precedes the focus constituent in wh-cleft (and it-cleft) sentences, the examples above clearly show that this is not the case with th-clefts. If the second copula (i.e. the copula directly preceding the focus constituent) is taken as main verb and question formation is applied, the ungrammatical sentence (7) is produced. Thus we see that the shared clause in th-cleft sentences and the presence of the spurious copula yield the unusual non-canonical surface structures exhibited.

Indefinite pronouns may precede the COMP-node in th-cleft sentences. That is, the surface structure complement NP in th-clefts is rewritten as either NP or as NP - $. Examples (8) and (9) clarify this statement:

(8) a. That's [[NP $$_S$$ what I never would have believed]] is that the
    earth is round.
b. That's [[NP $$_S$$ something] [$_S$ (that) I never would have be­
    lieved]] is that the earth is round.
(9) a. That's [[NP $$_S$$ what I hate]] is wet hair.
b. That's [[NP $$_S$$ something] [$_S$ (that) I hate]] is wet hair.

Lexical NPs, either definite or indefinite, may occur as the head of the complement as well, evoking the same rewrite rules as (8b) and (9b) above demonstrate.
That's a fact (that) I never would have believed, is that the earth is round.

That's a woman (that) I'd like to get to know, is Raquel Welch.

That's the one book (that) I would never read, is Aspects of the Theory of Syntax.

That's the German strategy (that) I never understood, was their invasion of Russia.

Akmajian (1970b:18 and 83n.) notes that pseudo-cleft sentences that begin with who are often less acceptable than those that begin with other wh-words. Consider sentences (14a) to (14f) (Akmajian's (2a) to (2f)).

(14) a. Who Nixon chose was Agnew.
    b. What Herman bought was that tarantula.
    c. Where he finally ended up was in Berkeley.
    d. When John arrived was at five o'clock.
    e. Why Fillmore sent Perry was to exploit the Japanese.
    f. How he did that was by using a decoder.

Interestingly, (14a) seems more acceptable in a th-cleft construction and all the sentences in (14) are certainly possible as th-clefts. That (15a) is better may be explained by the fact that the wh-clause appears in a complement position, however.

(15) a. That's who Nixon chose, was Agnew.
    b. That's what Herman bought, was that tarantula.
    c. That's where he finally ended up, was in Berkeley.
    d. That's when John arrived, was at five o'clock.
    e. That's why Fillmore sent Perry, was to exploit the Japanese.
    f. That's how he did that, was by using a decoder.

Furthermore, this, these and those may appear instead of that in certain contexts, although grammaticality judgments are ambivalent and my intuitions may be arguable. Consider sentences (16) and (17):

(16) a. This is what I hate to do, is my laundry.
    b. These are things (that) I absolutely despise, are washing the dishes and doing the laundry.

(17) a. That's what I hate to do, is my laundry.
    b. Those are things (that) I absolutely despise, are washing the dishes and doing the laundry.

I would suggest that the differences in meanings between (16) and (17) reflect variations in the immediacy of the situation. That is, in (16) the speaker is about to perform the chores and is expressing displeasure, while in (17) the statements are more general and need not occur within the context of the performance of these tasks. In other words, this and these are more restricted in occurrence, being directly dependent upon the context surrounding their utterance.

The factors which influence the tense of the copula also need mention. The phenomena exhibited in the examples are far from being clear at this
point. Consider:

(18) a. That's a cigarette (that) I used to hate, is Export A.
   b. That's a cigarette (that) I used to hate, was Export A.
(19) a. That's a road (that) I used to hate to drive on, is Crowchild Trail.
   b. That's a road (that) I used to hate to drive on, was Crowchild Trail.

The differences in interpretation are quite subtle. It seems to me that the (b) sentences above suggest that the item referred to by the focus constituent is no longer immediately present. That is, (18b) implies that Export A is no longer manufactured, while (18a) does not. Similarly, (19b) implies that the speaker no longer resides in Calgary, while (19a) does not.

These subtle semantic distinctions both in the use of the demonstrative pronouns and in the tense of the second copula may lend support to the grammaticality judgments which have been offered. Th-cleft sentences seem to add semantic material which cannot be expressed with wh-cleft (or, for that matter, it-cleft) sentences alone. The additional semantic yield to the speaker through the use of th-clefts could account for their presence in the grammar as an innovative construction.

To sum up this section, the significant fact that th-cleft sentences appear at the surface (roughly) in the form

\[ S_i \text{ Demonstrative Pronoun - be -} [S_j \text{ NP } S_i] - \text{ be } - \text{ NP } S_j \]

was demonstrated with the implication that this form accounts for the inherent strangeness and marginal acceptability of the construction. Moreover, evidence of the spurious nature of the second copula was presented. It was also shown that th-clefts may appear with any of the demonstrative pronouns in the subject NP position, depending upon the context of the situation. Finally, the problematic nature of the tense of the second copula was discussed.

3. Focus and Presupposition

Chomsky (1971) and Akmajian (1970a, b; 1973) both deal with the matter of sentence focus and presupposition in depth. In their terminology, the focus of a sentence is "a constituent of a sentence ... which contains the intonation center, i.e. the position of highest pitch and stress." (Akmajian 1970b:189-90) Akmajian goes on to add that focus refers both to emphatic stress in an utterance and to the normal peak stress point in an utterance. The presupposition of a sentence, on the other hand, is "a statement derived by replacing the focus of a sentence with an appropriate semantic variable." (ibid, p. 190) Furthermore, Akmajian states that "the general interpretation of the notion 'focus' is that portion of a sentence which is 'new', informative, 'interesting', and semantically prominent with respect to the surrounding material." (ibid., p. 192)
It seems that there are two basic intonation and stress patterns associated with th-cleft sentences, both of which can be considered marked patterns in the sense that in neither instance does the sentence follow the normal rules of stress assignment in English. Given that we are dealing with a focus construction here, this markedness is not unexpected.

Returning to a comparison of pseudo-cleft and th-cleft constructions, the two patterns are of some interest. In the first instance, the pseudo-cleft and the th-cleft intonation contours are almost identical. The only difference is that the th-cleft sentence has a short segment at the head of the first breath group. Consider sentences (20) and (21):

(20) a. What I won't do is go to the bar alone.
   b. That's what I won't do, is go to the bar alone.

(21) a. What I find utterly ridiculous is student protest of this nature.
   b. That's what I find utterly ridiculous, is student protest of this nature.

The foci of the sentences (20) and (21) are defined as those constituents which receive primary stress, namely bar in (20) and student protest in (21). Although not strictly identical, this pattern is analogous to the changes which occur in sentence stress after the there-insertion transformation is applied. The patterns associated with there-insertion are exhibited in (22) and (23):

(22) a. A fly is on the wall.
    b. There's a fly on the wall.

(23) a. A snake is crawling through the door.
    b. There's a snake crawling through the door.

The contrasts between the (a) and (b) sentences in (22) and (23) above are very similar to the contrasts which appear in sentences (20) and (21). The lack of stress on there in (22b) and (23b) reflects the fact that it has little or no semantic content. I suggest that the demonstrative pronoun in (20b) and (21b) behaves in a similar manner; that is, with its use, very little semantic content is added to the sentences in this instance. This assumption, if correct, lends support to the view that there is a rule of th-cleft formation similar to there-insertion (I will return to this in Section 4.).

The other form of sentence stress pattern is quite distinct from the first. Consider sentences (24) and (25):

(24) a. What I won't do is go to the bar alone.
    b. That's what I won't do, is go to the bar alone.

(25) a. What I find utterly ridiculous is student protest of this nature.
    b. This is what I find utterly ridiculous, is student protest of this nature.
My intuitions are not entirely clear, but it seems that in (24b) and (25b), the initial demonstrative pronoun serves as a constituent which adds emphasis to the sentence. I take this to indicate an addition of semantic content, albeit a minor one. Thus, we can see that the focus is split to include two coreferential constituents in these special cases.

More importantly, however, the differences exemplified in sentences (24) and (25) are significant in that further support is provided for the grammaticality judgments discussed in Section 2. In other words, the th-cleft sentences in (24b) and (25b) make subtle semantic distinctions which are not possible with the corresponding pseudo-cleft sentences in (24a) and (25a).

This section of the paper has discussed the sentence stress patterns associated with the foci and presuppositions of wh-cleft and the th-cleft sentences. I expect that there may be objections put forth to the analysis presented here, but I will, for the present, assume that it is basically correct. In Section 4, a proposal for the derivation of th-cleft sentences is presented which appears to account for a considerable proportion of the data.

4. Derivation of Th-Cleft Sentences

At one point, transformational grammarians (Chomsky 1970b:209-10; Akmaian 1970a, b) postulated the derivation for pseudo-clefts which is exhibited in (26):

(26) a. [NP[S-SCOMP[S I hate [NP wet hair]]]] be [PRED Δ ]
   b. [NP[S-SCOMP[S I hate [wh-]]]] be [NP wet hair]
   c. What I hate is wet hair.

This formulation provides a dummy node in the underlying structure. The focus constituent is extraposed to fill the dummy leaving behind a wh-"pro-form." Finally, wh- is fronted to the COMP node and agreement rules apply to the copula to form the surface structure in (26c).

It seems that this is quite a controversial analysis, particularly with reference to the revised extended standard theory. The important objections here are that the tensed-S, the specified subject, and the subjacency conditions are all violated in (26).

The tensed-S condition (also termed the propositional island constraint (PIC)) is formulated as follows:

(27) "... X ... [α ... Y ... ] ... X ..." (Chomsky 1977b:74; Chomsky's (11))

where α is a cyclic node and where "PIC asserts that no rule can 'involve' X and Y where α is a finite clause (tensed-S)." (ibid) This condition is
violated in (26) where *I hate wet hair* is Y and therefore cannot be involved in any movement rule which removes all or part of the clause.

Also applicable to (27) is the specified subject condition which states that "no rule can involve X and Y ... where a contains a subject distinct from Y and not controlled by X." (Chomsky 1977a:176) Once again, I is the specified subject within the embedded clause in (26a) and thus *wet hair* cannot be moved from its position in the clause.

Finally, sentence (26) violates the subjacency condition in its derivation. Chomsky (1977b) states the subjacency condition as follows:

(28) "A cyclic rule cannot move a phrase from position Y to position X (or conversely) in [the following structure]:

... X ... [α ... [β ... Y ... ] ... ] ... X ..., where α and β are cyclic nodes." (p. 73; Chomsky's (6))

Whether we take the cyclic nodes to be NP and 5, or NP and S, sentence (26) violates the condition. That is, the complement NP *wet hair* is moved rightward over all three node categories in the derivation, thus violating subjacency.

No clear non-problematic solution to this dilemma is apparent to me at present, at least in the case of pseudo-cLEFTs. The only solution which is reasonable is to base-generate all pseudo-cLEFT sentences, as Higgins (1973) and Halvorsen (1978) have suggested.

Another objection may be raised to the derivation demonstrated in (26) if one accepts the proper-binding constraint of the trace theory of movement rules. Freidin (1978) provides the following passage which outlines the basic notions here:

"(8) Proper Binding (PB)
Each bound anaphor αj in a phrase marker Pj must be
a. bound to some antecedent in Pj; and
b. c-commanded by its antecedent." (Freidin 1978:521)

where c-command is defined as follows:

"Node A c(onstituent)-commands a Node B if neither A nor B dominates the other and the first branching node which dominates A dominates B." (ibid. 521n)

In the derivation exhibited in (26), the NP *wet hair* does not meet the conditions of PB as outlined above. Although the NP does have an antecedent in the phrase marker (i.e. *what*), the constituent is not c-commanded by that antecedent. The NP *wet hair* actually appears at the surface in a higher sentence than is its antecedent. Hence the derivation which appears in (26) fails on several counts and must be rejected.

Following Fiengo's (1977:47) trace theory account of *there*-insertion,
a reasonable proposal for the derivation of th-cleft constructions can be formulated. Neither subjacency, PIC, nor the specified subject condition are violated in this formulation, but rightward movement does occur. The proper-binding constraint, however, is not violated. The derivation is demonstrated in (29), omitting some of the obvious steps:

(29) a. \([\text{NP wet hair}] \text{ is } [\text{NP} \text{S}\text{COMP}_s [\text{S I hate [+WH]}])]\]
b. \([\text{NP}_4 \text{t}] \text{ is } [\text{NP} \text{S}\text{COMP}_s [\text{S I hate [+WH]}]) [\text{NP}_4 \text{ wet hair}]\]
c. \([\text{NP}_4 \text{ that} ] \text{ is } [\text{NP} \text{S}\text{COMP}_s [\text{S I hate [+WH]})] [\text{NP}_4 \text{ wet hair}]\]
d. \([\text{NP}_4 \text{ that} ] \text{ is } [\text{NP} [\text{S[what]}_j \text{ I hate [t]}_j]] [\text{NP}_4 \text{ wet hair}]\]
e. That's what I hate -- wet hair.

A brief clarification of the steps in the derivation (29) is as follows: (1) the subject NP wet hair is extraposed to sentence-final position, leaving a trace [ t ] in (b)\(^g\); (2) the trace is covered by the demonstrative pronoun that in (c), thus avoiding the improper binding restriction; (3) [+WH] is fronted to the COMP node in S by wh-movement and leaves behind [ t ], in object position; and, finally, (4) the surface structure is realized by the (optional) contraction of that is to that's.\(^9\)

This derivation results in the surface structure (29e) which is considered acceptable by almost all speakers. To derive the final structure, an optional insertion transformation which may be called Copula Reiteration is proposed:

(30) That's what I hate, is wet hair.

\[\text{(31) SD: X - Dem Pron - be - NP - NP - Y} \]
\[1 \quad 2 \quad 3 \quad 4 \quad 5 \quad 6 \]
\[\text{SC: 1 \quad 2 \quad 3 \quad 4 + be 5 \quad 6} \]

The structural description of Copula Reiteration clearly violates the condition of minimal factorization proposed by Chomsky (1977a). This condition "requires that the SD cannot contain two successive categorial terms unless one or the other is satisfied by a factor changed by the rule." (p 172) I would argue, however, that the transformation (31) is a very low-level rule (i.e. one of Bach's "housekeeping rules") and, for some speakers, quite marginal. Hence, the minimal factorization condition is not necessarily a serious objection, but rather a minor obstacle to the formulation and application of the rule.\(^10\)

Note that in (29c), the trace is covered by the demonstrative pronoun, a move that has precedent in the trace theory account of there-insertion (see Fiengo 1977:47). The prediction which is made by the derivation in (29) is that th-cleft constructions are variants of right-dislocated constructions, which seems to be the case.
This analysis is still problematic, but, after months of travail with these sentences, seems to me to be the most useful account available.

Sentences which contain verbal complements as focus constituents are the most damaging to this analysis. Consider the underlying structure presented in (32):

\[
\begin{align*}
(32) \ &= \ \left[ \text{go to the bar alone} \right] \ \text{be} \ \left[ \text{NP} \ [\text{S} \ \text{COMP} \ [\text{I won't do } [+\text{WH}]]] \right]
\end{align*}
\]

The structure in (32) is problematic because we would expect to find the structure \( \text{NP} - \text{be} - \text{NP} \) underlyingly. This objection may be countered, however, by pointing out that the phrase \text{go to the bar alone} in (32) behaves like an NP in surface structure; that is, after extraposition, the phrase functions as a predicate nominal.

5. Conclusion

This paper has described an English construction which, to my knowledge, has not been discussed elsewhere. A reasonably well-motivated proposal was made for the syntactic account of the construction. Problems such as tense agreement patterns (sentences (18) and (19) and demonstrative-copula deletion (or insertion) (sentences (1), (14), and (15)) remain troublesome. It is hoped that further exploration into the nature of these constructions can provide properly motivated explanations for these objections.

Most importantly, however, a number of intersentential relationships are accounted for; in particular, the relationship between sentences which undergo Copula Reiteration and those which do not was established. A significant generalization concerning th-cleft formation and there-insertion (i.e. transformations that insert lexical items which have low semantic loads) was made, as well. It is important also that th-cleft constructions fill a semantic function which is not available through the use of other cleft constructions.

As a final note, it may be said that the derivation of th-cleft sentences, as it has been presented here, accords well with the notion of cognitive strategies in sentence production. That is, the proposed Copula Reiteration rule seems to fit a general theory of speech production in that factors such as short-term memory are taken into account. The notion appears promising, but requires much more exploration before anything definite may be stated.
Footnotes

*I wish to express my thanks to V. P. De Guzman and W. D. O'Grady who have been very helpful with their comments of the paper. William O'Grady, in particular, has always been available for discussion of difficult points in the analysis. Needless to say, any mistakes remaining are my own.

1The question of whether wh-cleft sentences are base-generated or derived is still apparently unsettled. For relevant discussion and proposals, see Chomsky (1970b:209-10) or Akmajian (1970a, b). Cf. Faraci (1971) or Nakada (1973) for the fundamental notions of the "embedded-question" analysis; Peters and Bach (1971) for an exposition of the "deletion" analysis; and Higgins (1973) or Halvorsen (1978) for arguments supporting a single source in the base for pseudo-cleft sentences.

2Some speakers find this sentence ungrammatical; that is, they find a pause preceding the focus constituent (the predicate NP of th-cleft sentences) acceptable, but a form of the copula in that position unacceptable. I will assume, however, that (lc) is acceptable within a certain context, on the grounds that (i) sentences of this type are documented (see J. Szarkowski 1970:15) and (ii) these constructions are heard often in conversation. Regardless, a juncture obligatorily occurs before the focus constituent which is marked with a comma.

3Th-clefts do not, at least, have a structure which corresponds in the same manner. The deep structure which is postulated for th-clefts in Section 4 is as close as one can come to such a correspondence, e.g. "That the earth is round is what I never would have believed."

4See Section 4 for a discussion and possible explanation for this fact.

5It should be noted that the relative pronoun is obligatorily deleted in (8a) and (9a). See Chomsky and Lasnik (1977:Section 1.B.) for relevant discussion.

6The valid objection may be raised that semantic distinctions have nothing to do with grammaticality judgments. I defend the use of grammaticality in this manner by pointing out that grammaticality is poorly understood within the current theory of grammar. My use of the term in this paper refers to the usefulness of the th-cleft; that is, since I am working with a construction which is highly unusual anyway, it seems that the best guess as to what a speaker might say must be presented. Thus, a "grammatical" sentence in this sense is one which might be uttered in conversation.

7But see Postal and Pullum (1978) for criticisms of trace theory.

8It should be added that a phonological juncture (pause) is obligatorily inserted at the surface in the space which precedes the moved NP.
The derivation of the sentences containing indefinites is similar, e.g.

(i) That's something (that) I hate -- wet hair.

The underlying structure for (i) would be as in (ii):

(ii) \[ \text{NP wet hair} \text{ is } \text{NP something [SCOMP [S I hate something]]} \]

Chomsky and Lasnik (1977) posit a transformation which violates the minimal factorization condition, i.e. their "It-insertion" rule (p. 449). This indicates that there is precedent for positing a rule which violates the condition.

References


