English Language and Linguistics 3.2: 335–352. © Cambridge University Press 1999.

Thing is constructions: the thing is, is what's the right analysis?[1]

DIANE MASSAM
University of Toronto
(Received 9 July 1998; revised 6 May 1999)

This paper examines an English construction which is very common in speech, but rare in writing. This construction is referred to as the *thing is* construction, as in: *The thing is, is we've got to be strong*. First, similarities between this construction and specificational pseudocleft constructions are noted, and it is proposed that the constructions are related. Then, five differences between the two construction types are discussed. An analysis is presented which accounts for the similarities and differences. Central to the analysis is the fact that *thing is* constructions contain 'theta-role by recognition' (Stowell, 1981), or appositive nouns (*claim*, *problem*, *feeling*, etc.), which allows the initial Wh-word to become optional. The paper also considers reduced *thing is* constructions such as: *Our kids are great on vacations, but when they come back, is, they need to play*, claiming that these, in a continuum along with clefts and *thing is* constructions, demonstrate the extreme case of a tendency for *be* to be used as a focus marker in spoken English.

1 Introduction

There is a construction in English which appears very often in spoken English, but extremely rarely in written English. Although it is robust and frequent, even linguistically sophisticated speakers are generally unaware of it, and it has received virtually no formal analysis.[2] Typical examples appear in (1).[3]

(1) (a) The problem is, is that we can't find the evidence.
(b) Part of our whole thing is, is that we've got to be strong.
(c) My feeling was, was that she doesn't have a professional hold on the situation.
(d) One of the realities is, is that we have hit the wall with respect to spending.
(e) The plan is, is that, if everything goes according to plan, is that the responsibility will be passed along to the departments.
(f) The hitch is, is that it seems to be occurring in the opposite direction.

[1] I would like to thank Jean Balcaen, Jack Chambers, Jun Mo Cho, Elizabeth Cowper, Jila Ghomeshi, Caroline Heycock, Jack Hoeksema, Alana Johns, Ellen Prince, Yves Roberge, Margaret Speas, Carolyn Smallwood, Bas Aarts, two anonymous reviewers, and members of the syntax project at the University of Toronto for their help with this work. I alone am responsible for any errors. This work has been supported by a Social Sciences and Humanities Research Council of Canada grant #410–97–0493.

[2] One paper that treats these constructions is McConvell (1988), which provides Australian, British, and American examples.

[3] Most original grammatical examples in this paper are taken from actual speech, unless specifically identified otherwise. Most examples are from semi-formal environments, such as university administrative meetings and class presentations. Some are taken from the media, in particular, radio news shows. Still others are taken from casual conversations. The ages of the speakers for whom I am aware of age, range from 5 to 59. All spoken examples collected by me were heard in Toronto, Ontario.

https://doi.org/10.1017/S136067439900026X Published online by Cambridge University Press

(g) The secret is, is for them not to find out. *(spoken by 5-year-old)*
(h) The only thing is, is I couldn't move down here because I don't drive.
(i) And my hunch is, is that there won't be any radical change.
(j) The important thing is, is we got to have lunch together.
(k) The fact of the matter is, is that the policy is totally unclear.

The normal first reaction to these sentences is that they contain a redundant instance of the verb *be*. In fact, some linguists and students have reacted to these sentences by considering them to be performance errors. However the frequency of the construction and the consistency of its form argue that it is a legitimate English construction susceptible to systematic analysis, as noted by McConvell (1988). Although it appears most commonly in spoken language, it can occasionally be found in writing. McConvell (1988) reports that instances appear in student essays, and Ellen Prince (p.c.) notes finding examples in electronic newsgroups. I have found one literary example, given below, where the context is that a real estate agent is asking when his client would like to visit a property.

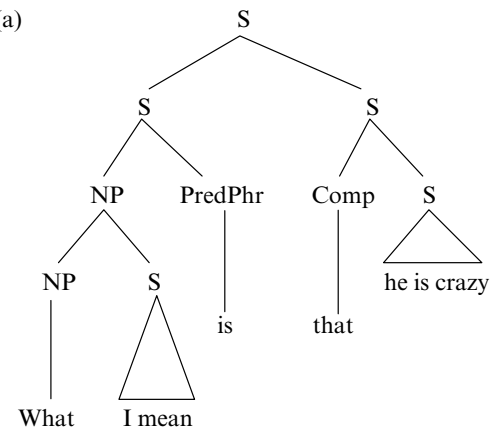
(2) Because, the thing of it is, is, there's another party interested, as well? (L. R. Wright, *Strangers Among Us*. Toronto: Doubleday Canada Limited 1996: 173) (The punctuation used is that of the original.)

I refer to the constructions as *thing is* constructions (henceforth T-i constructions) since they most commonly appear with the NP *the thing* as in (1h) and (1j) above. The question then is, what is the appropriate structural analysis of T-i constructions?

McConvell (1988) concludes that they are the result of the process of blending (Riley & Parker, 1986), which creates an amalgam of two possible structures.

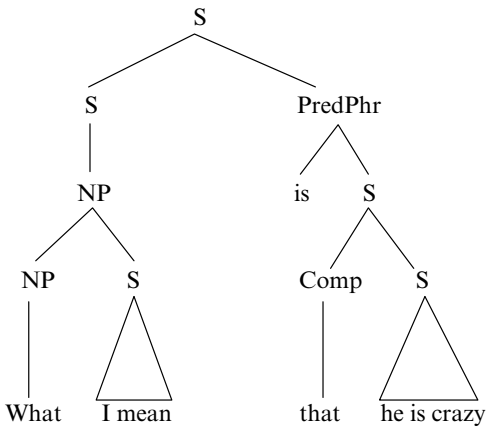
(3) (from McConvell, 1988)

(a)



https://doi.org/10.1017/S136067439900026X Published online by Cambridge University Press

(b)



McConvell states that for the resulting configuration, 'in the nature of blends, structure is somewhat indeterminate' (McConvell, 1988: 300). The theory of blending may or may not explain the origins of such structures, but to my thinking this analysis seems incompatible with the claim that these sentences are not performance errors. In a complete theory of grammar, each grammatical sentence should be assigned a unique syntactic structure, and structure cannot be indeterminate. With this in mind, this paper seeks to develop a syntactic analysis of these constructions.

First, the properties of these constructions are examined, taking as a starting point their superficial similarities to specificational pseudo-cleft constructions. The similarities and differences between these two construction types are observed. An additional construction, the reduced T-i construction, is exemplified which suggests that English is developing a presentential focus marker. Finally, an analysis is presented which accounts for the properties of the constructions.

2 T-i constructions and pseudocleft constructions

2.1 Background issues

Our first observation is that T-i constructions are similar in syntactic and intonational form to pseudocleft (PC) constructions with the verb *be*, such as those in (4).[4]

(4) (a) What Poirot$_i$ is, is important to himself$_i$.
(b) What Poirot$_i$ is, is important to him$_i$.
(c) Who Poirot is, is none of your business.

PC constructions have received attention in the linguistic literature for several decades (e.g. Akmajian, 1970; Bošković, 1997; Heycock & Kroch, to appear;

[4] Ignored in this analogy are pseudoclefts with verbs other than *be*, since, for reasons discussed below, T-i constructions always include the verb *be*. Note that another possibility for analysis is that T-i constructions are related to relative clauses such as: *The thing (which) is, is that she won't like it.* In this view there would be an assumed value, i.e. 'the thing which is important or relevant'. Given that the version with the overt Wh-word is ungrammatical, I have not pursued this possibility.

https://doi.org/10.1017/S136067439900026X Published online by Cambridge University Press

Higgins, 1979; Izvorski, 1998; Partee, 1986; Williams, 1983, among others). They have been divided into different types, depending on their syntactic and semantic characteristics. In the order given, (4) exemplifies the specificational type (a), the predicational type (b), and the indirect question type (c).[5]

Without delving too much into the complexities of pseudo-cleft constructions for the moment, we might attribute to them a syntactic structure in which the Wh-word binds a variable appearing after the first instance of *be* (henceforth *be1*).[6] Then, the entire Wh-clause preceding the second instance of *be* (henceforth *be2*), is the subject of the clause. Following Heycock (1994) I refer to the second part of the clause as the *counterweight*. I refer to the initial clause as the *setup*.[7]

(5) [[What$_i$ Ms. Trotwood is t_i] is a woman who hates donkeys]

This partial analysis explains the double *be* in the sentences in (4). The sentences in (1) might be analysed as PC constructions also as in (6), where the Wh-word just happens to be null. In this way the double *be* in (1) could be explained.

(6) [[OP$_i$ the problem is t_i] is that we can't find the evidence]

This analysis might be supported by the example in (7) which shows that there can be minimal pairs with and without an overt Wh-word in such constructions.

(7) I think what the whole thing is, is that he is trying to spearhead the whole campaign.

If T-i clauses are to be analyzed as a form of PC construction, we must determine which type of PC construction they are to be classed with.

We can eliminate the predicational type and the embedded question type in (4b,c), by noting that these sentence types differ from (4a) and from the T-i examples in (1), as exemplified by the following breakdown of paraphrases of the examples in (4a–c) and (6). The hatch sign indicates a lack of correspondence between the two sentences.[8]

(8) (a) What Poirot is, is important to himself.
(b) Poirot is important to himself.

[5] I base this on the classification of Higgins (1979). Some might argue that the term *pseudocleft* should be restricted so as not to apply to all types in (4).

[6] In particular, the various types of pseudoclefts require different analyses. While in the predicational type the Wh-clause might be best analysed as a headless relative, the embedded question Wh-phrase might receive a sentential analysis, and, as discussed most recently by Bošković (1997), the specificational type might require a more complex structure, given its particular characteristics. Others, such as Izvorski (1998) and Heycock & Kroch (to appear), argue that specificational pseudoclefts are a kind of equative sentence. The important point here is simply that the Wh-word in all cases will bind a post-*be1* variable.

[7] In PC literature the initial clause is usually referred to as the Wh-clause. This term does not work in this paper as I wish to include sentences such as those in (1) in the discussion, which precisely do *not* contain a Wh-word. I use setup, since to me it describes clearly the function of the initial clause. When I use the term Wh-clause, I am referring specifically only to PCs.

[8] There are additional arguments that T-i constructions are dissimilar from predicational pseudoclefts, as will be mentioned below.

https://doi.org/10.1017/S136067439900026X Published online by Cambridge University Press

(c) What Poirot is, is important to him.
(d) #Poirot is important to him.

(e) Who Poirot is, is none of your business.
(f) #Poirot is none of your business.

(g) The problem is, is that we can't find the evidence.
(h) The problem is that we can't find the evidence.

By a process of elimination, then, we are left to explore the possibility that T-i constructions are similar to specificational PC constructions.

2.2 Similarities between T-i and specificational PC constructions

Many syntactic characteristics of specificational PCs, noted in the literature, and reviewed by Bošković (1997), are shared by T-i clauses. For example, it is impossible to Wh-extract the counterweight. This is true whether or not an overt Wh-word appears in the T-i construction, as indicated by the optionality parentheses around *what* in (9b). (9a), taken from Bošković (1997), shows that Wh-movement of the counterweight is impossible with specificational PCs. (9b) shows that the same is true for T-i construction counterweights.[9]

(9) (a) *I wonder [how tall]$_i$ what John is, is [t]$_i$?
(b) *I wonder [what]$_i$ (what) the point is, is [t]$_i$?

In addition, as discussed in Hankamer (1974), Higgins (1979), and Williams (1983), it appears to be impossible for the setup clause to undergo A movement (unlike in predicational PCs: *What Mary is seems to be important to her*). This holds whether or not an overt Wh-word appears.

(10) (a) *What Steerforth is, seems t to be selfish.
(b) *(What) The point is, seems t to be that Emily will be late tomorrow.

A third property shared by specificational PCs and T-i constructions is that the counterweight is a barrier to movement, as discussed by Hankamer (1974) and Higgins (1979) and reviewed in Bošković (1997). Again, the facts are the same whether or not an overt Wh-word appears in the construction.[10]

(11) (a) *Who$_i$ do [you think [that [what Dora did] was confuse t$_i$]]?
(b) *Who$_i$ do [you think [that [(what) the point is], is that Dora confused t$_i$]]?

[9] It is impossible to set up a perfect illustration of the inability of the counterweight to undergo Wh-movement. Since subject/auxiliary inversion is not possible in T-i or PC constructions, main clause questions cannot be used. But, generally speaking, T-i constructions are ruled out in embedded position. I provide (9b) as the best Wh-movement construct I can devise. Note further that a test using topicalization is ruled out by the fact that a topic cannot be a focus (as pointed out by an anonymous reviewer). Thus, (i) (and perhaps (9b) also) is ungrammatical due to information content conflict.

(i) *[That Dora was in the boat]$_i$, (What) the point is, is [t]$_i$

[10] As with the Wh-movement discussed above, this fact must be placed in the context of the general infelicity of T-i constructions in embedded clauses. Note that in (11a) I use a PC with the verb *do* rather than *be*, since this provides the clearest example of the ungrammaticality of extraction.

https://doi.org/10.1017/S136067439900026X Published online by Cambridge University Press

These characteristics might indicate that the construction under discussion is similar in its syntactic properties to the specificational PC construction. However the PC construction facts are to be accounted for, the similarities can be accounted for by adopting an analysis of T-i constructions which considers them to be the same as PC constructions.[11]

Semantically too, T-i clauses and specificational PCs are similar, since in both the counterweight specifies the focused property left empty in the setup clause, and in both the setup is not referential. Higgins (1979) states that a specificational PC functions rather like a list in which the subject phrase constitutes the heading of the list and the predicate complement is an item on the list. Thus, according to Higgins, specificational PC clauses are interpreted as (12a). In (12b), an attested T-i construction shows that the two clause types are semantically similar.[12]

(12) (a) John is this: silly. (Higgins, 1979:8)
(b) The thing is, is this: we can't possibly raise the money.

According to Akmajian (1970), the Wh-clause of the clefted sentence contains a semantic variable (bound by the Wh-word) and this variable is specified by the focused postcopular item. Another approach is presented in Izvorski (1998) and in Heycock & Kroch (to appear), where PC sentences are interpreted as equatives, where equation is interpreted as a function assigning values to variables. Thus specificational PCs contain both a focus and a specificational (or equational) meaning.

From the above discussion, it might be concluded that T-i clauses are simply PC constructions with a null Wh-word. Whatever account can be devised to explain PCs, then, could be used to explain T-i constructions. However, in many ways, T-i clauses are unlike PCs, and these differences must also be accounted for. Below I outline several differences between the two constructions.

2.3 Differences between T-i and PC constructions

I will first consider one difference between the two types of constructions, only to conclude that it is not a true difference. Then, five true differences are reviewed.

Higgins (1979) discusses several unusual characteristics of specificational pseudo-clefts. The most often discussed of these is connectivity. This refers to the fact that in many ways the two separate phrases of the sentence behave as if they were one. For example, the subject of the Wh-clause can bind a reflexive in the counterweight, in spite of the fact that it does not c-command into the counterweight. (This is exemplified in (4a) above, repeated below.) Thus, for binding, the two parts of the

[11] The main point of this paper is to show that T-i constructions can be analyzed as PC constructions, not to argue for a particular analysis of PC constructions.

[12] The difference between these two sentences is due to the fact that (12a) is provided as a sort of Logical Form representation, hence the Wh-clause is simplified to a noun phrase (*John*) whereas the second clause is a spoken example. The similarity is that in both, the counterweight is set up to be the specification of the list.

https://doi.org/10.1017/S136067439900026X Published online by Cambridge University Press

sentence act as if they were connected, since normally an anaphor must have a c-commanding antecedent within its minimal sentence.[13] The opposite scenario holds for predicational pseudoclefts, in which pronouns appear rather than reflexives, as seen in (4b), repeated below.

(4) (a) What Poirot$_i$ is, is important to himself$_i$.
(b) What Poirot$_i$ is, is important to him$_i$.

T-i clauses behave differently from both types of PCs in this way, since in T-i clauses, no coreference is possible at all. The reason why connected reflexivization is not possible is that the binding domain for a counterweight reflexive would always be no larger than the counterweight itself, since we are dealing with sentential counterweights. Pronominal coreference, too, is ruled out, in a way that seems related to the *i-within-i* constraint of Chomsky (1981), as seen in (13).

(13) (a) *(What) The point$_i$ is, is [to make it$_i$ clearly]$_i$
(b) *(What) The point$_i$ is, is [its$_i$ clarification will help everybody]$_i$

The *i-within-i* constraint states that a constituent may not be coreferential with a subpart of itself, thus ruling out clauses like (14).

(14) [a picture of [itself]$_i$]$_i$

In (13), since *the point* is in an identity relation (as will be discussed below) with the counterweight sentence, it cannot also hold an identity relation with a subpart of the counterweight.

Thus, the most famous of specificational and predicational PC characteristics, that of connectivity, does not appear to hold in T-i cases. This is true whether or not an overt Wh-clause appears in the sentence, as is shown in (13) by the optionality parentheses around the Wh-word.

Before determining that the above facts disqualify T-i constructions from receiving a specificational PC treatment, consider the following. It has been observed that bound pronoun relations and negative polarity relations, as well as binding relations, treat the specificational PC as connected, i.e. as if the entire sentence consisted of only one sentential clause, as illustrated in (15). In these sentences, under normal circumstances, the italicized items would require that the underlined items c-command them within a minimal clause. This does not appear to be the case in (15), but the sentences are nonetheless grammatical.

(15) (a) What everyone proved was *his or her* own theory.
(b) What Murdstone didn't buy was *any* pictures of Copperfield.

It is not entirely straightforward to determine if these facts hold for T-i constructions. The first, concerning bound pronoun relations, involves the same *i-within-i* problem noted above, as seen in (16). Furthermore, quantifiers are not possible at all in the setup clause (cf. Iatridou & Varlokosta, 1997).

[13] Bošković (1997) points out that the opposite results obtain for binding of traces, but this effect cannot be tested in T-i constructions, since the relevant constructions are impossible for other reasons.

https://doi.org/10.1017/S136067439900026X Published online by Cambridge University Press

(16) *[Every point]$_i$ is, is [that [it]$_i$s own proof is impossible]$_i$.

The negative polarity facts are also difficult to test since I have heard no naturally occurring example of a T-i construction with a negative setup.[14] (17a) contains a made-up example. It is difficult to access intuitions on these constructions, but (17a) is clearly more acceptable than (17b) (with the nonstressed reading for *anyone*).[15]

(17) (a) ?My doubt is, is that anyone will run for mayor.
(b) *My point is, is that anyone will run for mayor.

Given the (near-)acceptability of (17a), I consider that connectivity holds in T-i constructions, as it does in specificational PCs, and that other factors rule connectivity effects out in cases of pronominal or reflexive binding. Connectivity can thus be considered to be a shared characteristic of the two constructions.

We now turn to some real differences between PCs and T-i constructions. The first, most striking difference between PCs and T-i constructions is so fundamental that it throws in doubt the idea that the two constructions are to be analytically related at all. This is the simple fact that the Wh-element in all clefts other than the type exemplified in (1) is completely obligatory. Thus, the sentence in (18b), related to (18a), is impossible.

(18) (a) What Poirot is, is a pompous fool.
(b) *Poirot is, is a pompous fool.

If we accept a straightforward PC analysis of the sentences in (1), then we must explain why the Wh-word can be absent just in particular circumstances. Before this can be done, it is necessary to characterize the circumstances wherein the Wh-word can be absent. These are as follows: in T-i sentences, the relation between the NP in the setup (*thing*, *claim*, *problem*, etc.) and the counterweight clause must be one of 'theta role by recognition' (Stowell, 1981). This means that the class of nouns permitted in the first part of a T-i construction is the same class of nouns which appears with a direct sentential complement, as in (19).[16]

[14] It is interesting to note that T-i constructions seem to become impossible when directly negated.

(i) The point *never is*/*is not* that people are being exploited.
(ii) *The point *never is*/*is not*, is that people are being exploited.

I assume this has to do with the apposition relations in these sentences, as discussed below in the text. (But see example (29i) below in the text.)

[15] An anonymous reviewer argues that sentences like (17a) are ungrammatical independently of the negative polarity item, thus ruling out (i).

(i) My doubt is that John will come.

Others have found more or less difficulty with accepting various nouns with *that*-clause complements, for example, one person noted difficulty with *problem* as in (ii), and with *issue*. I find them acceptable, though some are more clearly so than others.

(ii) the problem that he doesn't know what he is doing . . .

[16] Jack Hoeksema notes (pc) that this class of nouns is also the most common class which appears in determinerless nominal phrases, as in (38b,c) in the text.

https://doi.org/10.1017/S136067439900026X Published online by Cambridge University Press

(19) (a) the claim that the prime minister left early
(b) the reality that we have no money left
(c) the feeling that she didn't have a professional hold on the situation

Interestingly, the most common noun in *thing is* constructions appears to violate this generalization, since (20) might be considered an ungrammatical NP.

(20) the thing that Agnes cares about David

This is not the case, however, in spoken English where *thing* actually appears frequently with a direct sentential complement, as in the real-speech examples in (21).[17]

(21) (a) We were talking about this thing that everybody has to wear brown all of a sudden.
(b) It's like he's got this thing that he just doesn't want to be with me anymore.

Stowell develops the notion of 'theta role by recognition' to account for why some sentences can appear directly after a noun with no intervening case assigning preposition, as in (19). That such sentences can appear without case is a problem for Stowell's view that case is a prerequisite for thematic role interpretation, an issue which need not concern us here. What is of interest to us, is that his solution to this problem is to claim that such sentences do not receive true thematic roles from a predicate in the normal way. (Hence, he argues, they do not require case for thematic role interpretation.) Instead, he claims, there is an inherent relation of apposition or identity between the noun and the sentence. We refer to this as an apposition relation, and to the relevant class of nouns as apposition nouns.[18]

The stipulation that T-i constructions must include an inherent apposition relation allows us to derive the fact that the initial noun is always one which inherently licenses an apposition relation (*point*, *thing*, *claim*, *belief*, *problem*, etc.) as seen in the examples in (1) and elsewhere in the paper. It also explains why the verb in the sentence must be the verb *to be*, since the apposition relation allows for no other intervening verb. Finally, it explains why the counterweight material must be a clause, since the content of an inherent apposition noun is always a proposition. That these are indeed three facts is illustrated below.

(22) (a) *The desk is, is mine.
(b) *The point is, bothers me.
(c) *The answer is, is number three.

We thus see that a null or absent Wh-word is possible in a PC construction just in

[17] In fact, the use of the word *thing* in English is frequent in focus and emphasis situations, as in the following examples.

(i) This was the thing, I just didn't want to go fast.
(ii) That's been a real thing this year in our family.
(iii) You know that thing, where they look at you like this. [making a face]

[18] This could be seen as analogous to the view (cf. Déchaine, 1993; Rapoport, 1987, among others) that the relation between the two NPs in an equational sentence also does not involve any theta role assignment.

https://doi.org/10.1017/S136067439900026X Published online by Cambridge University Press

case the relation is one of theta role by recognition, or apposition. This holds as a stipulation now, but it must be explained, if a PC-based analysis is to be developed.

There is a second crucial difference between T-i and PC constructions, namely that in the former it is possible to have no verb *be* in the counterweight, as seen below.

(23) (a) The problem is, we can't find the evidence.
(b) The thing is, we got to have lunch together.

This is not the case for PC constructions.

(24) (a) *What I saw, an apple
(b) *What she is, a good kid
(c) *What he said, that the schoolboards were in a bind.

It must be noted that (23) differs from (25) below in that in (23) it is *be2* which is not present, whereas in (25) it is *be1* which is lacking.

(25) The point is that she doesn't want to run for mayor anymore.

This conclusion is based on the intonation patterns of these sentences, which is reflected in the use of commas in (23) as compared with no comma in (25). (See McConvell, 1988 for diagrams of the intonation contours involved.) In (23), there is high/falling/rising intonation on *is*, followed by a pause, then continuation. In (25), there is no varying intonation, neither on *point*, nor on *is*. Instead, there is simply a slight pause after *point*. Thus minimal pairs like in (3) exist, with members distinguished only by intonation, but nonetheless very clearly distinguished. Interestingly, with the noun *thing*, the pattern in (25) is unusual, perhaps following from the fact noted in footnote 17 that this apposition word is used mainly in cases of focus.

There is a third difference between T-i and PC constructions. Bošković (1997) points out that tag questions with specificational PCs are better if the tag is related directly to the NP and the verb within the Wh-clause, rather than to the ostensible main verb (*be2*).

(26) (a) *What Steerforth is, is selfish, isn't it?
(b) ??What Steerforth is, is selfish, isn't he?

The opposite scenario holds with T-i constructions in that they permit the tag to relate to material in the counterweight more readily than to the material in the setup. Here too, however, the tag does not appear to relate to the ostensible main verb (*be2*). I follow Bošković's direction in claiming the first sentences in (26) and (27) to be ungrammatical, though I feel that they are not fully so. Note, however, that the relative judgements hold.

(27) (a) *The point is, is that you want to run the meeting, isn't it?
(b) ??The point is, is that you want to run the meeting, don't you?

There is a fourth difference between T-i and PC constructions. As also observed by McConvell (1988), there exist what I term *reduced T-i constructions*, which involve the elimination of the apposition noun. Some examples are seen in (28)–(30).

https://doi.org/10.1017/S136067439900026X Published online by Cambridge University Press

Such data are even more restricted to spoken language than the examples in (1), if this is possible, and they are not so commonly heard as the regular T-i construction, which is rampant. However, the data here have been collected in natural speech environments, and such examples are not extremely rare.

(28) (a) Our kids are great on vacations, but when they come back, is they need to play.
(b) You can still account for productivity, is you just need a separate level.
(c) She was telling me, is they have to eat with the kids.
(d) I realized when he said that, is he just doesn't care.
(e) For people who know me, is I don't like confrontations or conflicts.

(29) (a) I think we have to do this, at least logically, is we have to go back to the way it was.
(b) Can I simply say this, is that the parliamentary process is a difficult one. (from McConvell, 1988)
(c) You know what, is we've got kids ages 7 and 10 and they've grown out of those videos.
(d) That's just it, is that we shouldn't have been there in the first place.
(e) I made the point once before, is that we have to *work* on this committee. (from McConvell, 1988)
(f) The fact remains, is the people's living standards are being cut. (from McConvell, 1988)
(g) You see that again and again, is that phonologists come up with theories for English which just don't work for Tiberian Hebrew.
(h) A couple of things are interesting, is they say that people aren't going to buy generators even though they all claim they're going to.
(i) That can't be a very welcome outcome, is that rates will now rise. (from McConvell, 1988)

(30) (a) That was what I was talking about, is that they do already do that. (from McConvell, 1988)
(b) That's the other thing I wanted to say about this, is that we should never have agreed to the conditions.
(c) Maybe that's why we're self-reliant, is we don't expect anything from the government.
(d) And really, that is one essential point, is that they have been running a losing campaign.
(e) Yes, which is the problem with the phonology paper, is that I was starting from scratch.
(f) I want to recommend something that might help, is that you must say to them when you are available before the committee is struck.

I have divided these data into three groups. The first, (28), I consider to be the clearest examples of the phenomenon under discussion. The second and third might be argued to be performance errors, although I do not believe that they are. In each example in (29), an apposition type of noun might be argued to be present if we consider deictics such as *this*, *what*, *it*, and *that* to be semantically similar to apposition nouns in that their contents can always be spelled out by a proposition. In (30) it might be argued that the speaker has forgotten that he or she began with

https://doi.org/10.1017/S136067439900026X Published online by Cambridge University Press

that's or *that is*, *which is*, or *I want to recommend*. Note that if we delete the initial *that's*, etc. we end up with a regular copular sentence or PC in (30a–d). The same is true for (30e,f), where deletion of *which is* or *I want to recommend* results in a regular copular sentence. In spite of these possibilities, like McConvell, I consider that these are not performance errors but exist as a well-formed speech construction, here termed the *reduced T-i* construction. Below, I suggest that *be* in such sentences is a pure focus marker.

These sentences, if related to those in (1), are problematic for a cleft analysis of T-i sentences in that they show that it is possible to have not only a null Wh-word, but also a null subject in the setup. Clearly this is not an option for other PC constructions, as seen below.

(31) (a) *Said, was that Uriah was going to be late.
(b) *Is, is a good kid.

A fifth difference is that it is possible to reverse a PC construction, but not a T-i construction, as seen below.

(32) (a) A fool is what he is.
(b) *That Peggoty is energetic, is the point is.

So far, then, we have outlined some similarities between T-i and PC constructions, which might be accounted for by adopting roughly the same syntactic analysis for the two construction types. We have also, however, pointed out five major differences between the two construction types which must be accounted for. In the next section, we will develop an analysis which allows both similarities and differences to be accounted for.

3 Focus and specification

In this section we will analyze the properties of the T-i constructions in a way which allows us to account for the similarities and differences between these constructions and PC constructions. The analysis depends on the two aspects of the construction isolated by Akmajian (1970): focus and specification.

In both PC and T-i constructions, let us posit an INFL head, which contains a Focus feature along with inflectional features and *be2*, as seen in (33).[19] INFL, which contains the focus/copular head, requires that some constituent appear in its specifier position due to the Extended Projection Principle (cf. Chomsky, 1995). It is the setup clause that satisfies this requirement. The material to the right of the focus/copular head in (33) is the focused predicate.

(33) [[What$_i$ Mary is t$_i$] is [a fool]]
[F]

[19] The assumption here is that *be* acts as support for [F]. In regular clefts it acts simultaneously as a verb and a focus marker, while in T-i and reduced T-i constructions, it can be simply a focus marker. An alternative would be to follow Déchaine (1993) for the clefts, and consider that [F] acts like her other functional heads and c-selects a VP with *be* as its head.

https://doi.org/10.1017/S136067439900026X Published online by Cambridge University Press

Following Bošković (1997), I assume that the counterweight moves covertly (i.e. at the level of Logical Form, or LF) into the setup clause, replacing the Wh-word.[20] I assume further that it then reconstructs back to the position of the variable (t_i in (33)). This movement derives the connectedness effects, since the clause is now unified at the level of interpretation, as shown in (34), and, as outlined in Bošković (1997), it also explains the syntactic constraints on movement discussed above.

(34) Mary is a fool is
[F]

Let us now turn to the analysis of T-i constructions. The claim is that T-i constructions are virtually identical to specificational PCs, except that rather than a Wh-word and its trace, the setup includes simply an empty variable (thus differing from (6)). In this case also, the counterweight, here a sentence rather than a predicate, will raise to the setup covertly, adjoining to the setup as shown in (35b), then reconstructing to (35c).

(35) (a) [[the thing is ec_i] is [that I like you]]
[F]
(b) [[that I like you]$_i$ [the thing is ec_i] is t_i
[F]
(c) [t_i the thing is that I like you] is t_i
[F]

By adopting a PC account of T-i constructions (here, that of Bošković, 1997) we can account for the intonational, semantic and syntactic similarities between the two construction types. We have yet to deal with two issues: the syntactic licensing of the empty category in T-i constructions, and the five main differences between the two construction types.

First, we deal with the licensing of the empty category in the T-i constructions, which allows for the absence of the Wh-word. How is the empty category in (35a) licensed? Recall Akmajian's observation that there is a specificational relation between the two clauses. This relation requires a variable in the setup which, according to Bošković, is ultimately replaced by the counterweight at LF. This variable, however, must be syntactically licensed independently of the counterweight. In the case of regular PC constructions, the variable in the setup is licensed by means of Wh-movement. In the case of apposition nouns, however, as discussed above, we can posit an inherent variable-binder within the apposition noun itself. This can directly bind a variable empty category in the setup clause, hence no Wh-word is required, although it may be present, as in (7).

This analysis also extends easily to explain the reduced T-i examples in (28)–(30), which receive the analysis below.

(36) is, [my kids have grown out of . . .]
[F]

[20] There are many analyses of PC constructions in the literature. I adopt Bošković (1997) here because it is the most recent paper which presents an explicit *syntactic* account of the PC properties.

https://doi.org/10.1017/S136067439900026X Published online by Cambridge University Press

For speakers allowing such sentences, the specifier of the focus/copular head need not be filled, with the result that T-i constructions for these speakers do not require a setup phrase. In not requiring a subject, *be* in such sentences has virtually completed the shift from being a copular verb with a focus feature appearing in INFL (to which the Extended Projection Principle applies) to being a pure focus marker in a Focus head (to which the Extended Projection Principle does not apply). Thus, English, like many languages, has developed a focus marker, i.e. a particle which attaches to a focused clause, from a copula. Other examples include Mandarin Chinese (cf. Zhang, 1997), and Haitian Creole (cf. Manfredi, 1993). Givón (1989) states that if focus constructions derive historically from the simplification of clefts or pseudoclefts, then one surviving relic of such simplification is the cleft-marking copular verb itself. This derivation appears to be occurring in spoken English.[21]

Along these same lines, when *be2* serves more of a focus particle role, it loses its matrix verbal status. This explains why tag questions do not treat *be2* as the main verb, neither in specificational PCs nor in T-i constructions, as we saw above. Remaining to be explained is why tags target *be1* in specificational PCs and the counterweight verb in T-i constructions. This is explained if in both cases it is the LF-connected clause that is targeted by the tag. The LF-connected clause for each construction type would be (37a) and (37b), respectively, where *be2* is deleted, as irrelevant.

(37) (a) Clara is a good kid [F]
(b) The point is that we can't find the evidence [F]

In the first example, a PC construction, there is only one subject and one true verb for the tag formation to target. In the second example, of a T-i construction, there are two, and it is the verb *find* in (37b) that is the more salient in terms of communicative content, as McConvell points out. In essence, the apposition noun is there only to license the focus construction, which, when it is parasitically based on the copular construction, requires a noun phrase in the specifier position to satisfy the Extended Projection Principle.

Indeed, at times when the setup in T-i constructions is present, it can be seen to act as a focusing adverbial rather than as the specifier and Extended Projection Principle subject of a focus/copular verb. This is seen in the examples below, where the semantically matrix status of the counterweight is evident, in all cases by the absence of *be2*, in (38a) by the embedded form of *be1*, and in (38b) and (38c) by the ungrammaticality of a complementizer. [22]

(38) (a) The thing being, I'll probably test positive.
(b) Thing is, (*that) he can't quit the committee right now.
(c) Problem is, (*that) they won't talk to the mayor about it.

[21] I am indebted to Jun Mo Cho and Niina Zhang for discussion of these points.
[22] Hoeksema (1999), in analyzing sentences such as (38b) and (38c), focuses on the fact that no determiner appears before the apposition noun.

https://doi.org/10.1017/S136067439900026X Published online by Cambridge University Press

The adverbial nature of the setup clause is also commented on by McConvell, who notes that in these and related constructions, 'the force of the clause which is formally subordinate is in fact that of the prime assertion, and the function of the clause which is formally main is to add information on the epistemological status of the prime assertion' (McConvell, 1988: 303).

Related to this is the fact, discussed in section 2, that *be2* is optional in T-i constructions, though obligatory in PCs. Again, in the former constructions, *be2* serves principally as an optionally overt reflection of the focus feature, and not as a verb. In the latter constructions, though, there is no other verb in the counterweight, and *be2* is also functioning as a verb in INFL, hence *be2* cannot be deleted.

The claim that *be2* is at least in part a focus marker also accounts for other observations which have been made about cleft sentences. In specificational PCs, *be2* can be only *is* or *was*, and cannot appear in any other form, an unusual restriction on a seemingly matrix verb. This requirement also holds in T-i constructions. Examples such as (39) have been attested, where the form of *be1* varies, but none where the form of *be2* varies beyond *is* or *was*. (An example of *was* as *be2* appears in (1c).)

(39) (a) The thing was, is that she would have been fine anyway.
(b) The thing was, is that we had no control over the situation.
(c) I suppose the other aspect was, is who's gonna be elected mayor of the megacity?
(d) The problem being, is that I'm probably going to test positive.
(e) The point of it being, is that there was no involvement.
(f) The cruel facts of life are, is that not every person who teaches Art is a good artist himself. (from McConvell, 1988)

Bošković (1997) also states that modals and adverbs such as *probably* are not possible in specificational PCs, and modals and adverbs are also not found in T-i constructions, as seen in (40). (Note the judgements below apply only to the specificational reading.)

(40) (a) *What Uriah is, can be important.
(b) *What Uriah is, is probably important
(c) *The thing is, can be that Dora will be late tomorrow.
(d) *The thing is, is probably that Dora will be late tomorrow.

Bošković accounts for these facts by claiming that there is no INFL in these sentences, only a VP. He is then forced to claim that the morphological agreement and the possibility for a past or present form is due to default and/or agreement specifications. Given the discussion in this paper, we can pin the near uniformity of the form of *be2* to its role as a focus feature bearer in a head position which may then not contain certain other functional features usually associated with INFL.

A final difference remains to be explained. As noted above in (32), it is possible to reverse a PC construction, but not a T-i construction. Bošković (1997) claims that the ability to reverse a specificational PC, as opposed to a predicational PC (seen in (41)), is due to the fact that the Wh-clause and the counterweight in a specificational

https://doi.org/10.1017/S136067439900026X Published online by Cambridge University Press

clause do not stand in a predicate–argument relation, unlike *What David is* and *important to him* in (41a) [23]

(41) (a) What David is is important to him
(b) *Important to him is what David is.

Since we are claiming that T-i constructions are like specificational PCs, and not like predicational PCs, it might be expected that they should be reversible, but they are not.

(42) *That we can't find evidence is the point is.

A clear solution to this problem will not be presented here. One possibility is that reverse constructions such as (32a) involve movement of the counterweight to the specifier of CP. It seems correct that T-i constructions do not have CPs, since they do not appear with complementizers, nor in embedded (nonbridge) contexts. In addition, given the focus marker analysis of the copular verb in these constructions, it is a natural step to assume that the FocP replaces the CP as the top extended sentential projection. If there is no CP, it is not possible for the *that*-clause to undergo topicalization to this position, hence the reverse construction will be ruled out. But intuition tells us that the problem with (32b) has more to do with the inability of the phrase *the point is* to be a postcopular constituent, rather than with the fronted material. If this intuition is correct, the ungrammaticality might be due to the fact that *be* cannot be followed by an IP, assuming that *the thing is* is an IP. This would be consistent with the subcategorization facts of the verb *be* in other contexts.

4 Conclusion

In this paper we have examined the syntactic properties of the English *thing is* construction such as: *The problem is, is that we can't find the evidence*. The proposed analysis is based on similarities between these constructions and specificational pseudocleft constructions. But there are also many differences between the two construction types. An analysis was developed which accounts for both similarities and differences. Central to the analysis is the fact that *thing is* constructions contain 'theta-role by recognition' or appositive nouns (as discussed by Stowell, 1981, such as *claim*, *problem*, *feeling*, etc.), which allow the initial Wh-word to become optional. The paper also examined reduced *thing is* constructions such as *Our kids are great on vacations, but when they come back, is they need to play*, claiming that these, in a continuum along with clefts and *thing is* constructions, demonstrate the extreme case of a tendency for *be* to be used as a focus marker in spoken English.

[23] Many linguists, such as Déchaine (1993), Heggie (1988), Moro (1997), and Williams (1983), explore the possibility that the Wh-phrase of a specificational PC is the fronted predicate, while the counterweight is the subject.

https://doi.org/10.1017/S136067439900026X Published online by Cambridge University Press

Author's address:
Department of Linguistics
University of Toronto
Toronto, Ontario
Canada
M5S 3H1
dmassam@chass.utoronto.ca

References

Akmajian, A. (1970). Aspects of the grammar of focus in English. PhD dissertation, MIT.
Bošković, Z. (1997). Pseudoclefts. *Studia Linguistica* **51**: 235–77.
Chomsky, N. (1981). *Lectures on government and binding*. Dordrecht: Foris Publications.
Chomsky, N. (1995). *The Minimalist Program*. Cambridge, MA: MIT Press.
Déchaine, R.-M. (1993). Predicates across categories. PhD dissertation, University of Massachusetts at Amherst: GLSA Publications.
Givón, T. (1989). *Syntax: a functional-typological introduction 2*. Amsterdam: John Benjamins Publishing Company.
Hankamer, J. (1974). On the non-cyclic nature of wh-clefting. In La Galy, M., R. Fox & A. Bruck (eds.), *Papers from the tenth regional meeting of the Chicago Linguistic Society*. Chicago: Chicago Linguistic Society. 221–33.
Heggie, L. (1988). The syntax of copular structures. PhD dissertation, University of Southern California.
Heycock, C. (1994). *Layers of predication: the non-lexical syntax of clauses*. New York: Garland Publishing.
Heycock, C. & A. Kroch (to appear). Pseudocleft connectedness: implications for the LF Interface level. To appear in *Linguistic Inquiry*.
Higgins, R. (1979). *The pseudo-cleft construction in English.* New York: Garland Publishing (original thesis date: 1973).
Hoeksema, J. (1999). Obligatorily inverted bare noun predicates. Paper presented at the Linguistics in the Netherlands meeting, Utrecht.
Iatridou, S. & S. Varlokosta (1997). A cross-linguistic perspective on pseudoclefts. In Kusumoto, K. (ed.), *Proceedings of NELS* 26. University of Massachusetts at Amherst: GLSA Publications. 117–31.
Izvorski, R. (1998). Specificational pseudoclefts have equative syntax. Talk presented at NELS 29. University of Delaware.
Manfredi, V. (1993). Verb focus in the typology of Kwa/Kru and Haitian. In Byrne, F. & D. Winford (eds.), *Focus and grammatical relations in Creole languages*. Amsterdam: John Benjamins. 3–49.
McConvell, P. (1988.) To be or double be? Current changes in the English copula. *Australian Journal of Linguistics* **8**: 287–305.
Moro, A. (1997). *The raising of predicates: predicative noun phrases and the theory of clause structure.* Cambridge: Cambridge University Press.
Partee, B. (1986). Ambiguous pseudo-clefts with unambiguous *be*. In Berman, S., J.-W. Choe & J. McDonough (eds.), *Proceedings of NELS 16,* University of Massachusetts at Amherst: GLSA Publications. 354–66.
Rapoport, T. (1987). Copular, nominal and small clauses: a study of Israeli Hebrew. PhD dissertation, MIT.

https://doi.org/10.1017/S136067439900026X Published online by Cambridge University Press

Riley, K. & F. Parker (1986). Anomalous prepositions in relative clauses. *American Speech* **61**: 291–306.
Stowell, T. (1981). Origins of Phrase Structure. PhD dissertation, MIT.
Williams, E. (1983). Semantic and syntactic categories. *Linguistics and Philosophy* **6**: 423–46.
Zhang, N. (1997). Syntactic dependencies in Mandarin Chinese. PhD dissertation, University of Toronto.

https://doi.org/10.1017/S136067439900026X Published online by Cambridge University Press