

Clefting and extraposition in English¹

Andreea S. Calude
University of Auckland

Abstract

Clefts and extraposition have structural and functional similarities, yet they have been mostly treated separately. An investigation of the Wellington Corpus of Spoken New Zealand English reveals cases where the two constructions appear confounded and difficult to disentangle. The present paper argues that they can indeed be differentiated, and provides a test which can be used to this end. This work offers contributions of a practical nature, in that it provides researchers with an objective criterion for distinguishing between clefts and extraposition, and also in a more theoretical sense, in its focus on two important strategies used in information packaging and discourse organization.

1 The problem of disambiguating between IT-clefts and extraposed clauses

The work detailed in this paper arises from a practical problem encountered when attempting to analyze the Wellington Corpus of Spoken New Zealand English² (henceforth WSC) for cleft constructions, in particular for IT-clefts. The occurrence of IT-clefts was investigated in a subpart of the corpus, namely in 193,000 words of spontaneous, face-to-face conversation data, which was found to contain 150 IT-clefts (this equates to roughly just over 75 constructions for every 100,000 words). Consider the two examples³ given in (A) and (B):

- (A) DS well margaret started wanting to wear makeup and so on when
she was onl <unclear word⁴>
MK yeah totally different type of child
DS yeah
MK or <vocalization> woman now of course
DS yeah young women now for sure
MK <laughs>

- DS oh **it's quite good** <drawls> **to to hear that faith's actually experimenting with a few** <long pause> **different things**
- MK <drawls> yes well raelene was she giggled <laughs>
- DS yes i'm sure she did <clears throat> (WSC#DPC002)
- (B) CH well they never got close to me but they were they're real cowboys and they're idiots you know and they're going on about this and that like and they said they wanted an under the table job and i said okay so that's <pause> w we agreed on the price <latch>
- BT should've
- DN mm
- AL you haven't paid them anything
- CH i haven't paid them a cent and they're meant to come back <pause> well they came back when i was on holiday i left this really clear message saying do not come back till i get back from holiday <pause> **it's about six weeks ago they did the job**
- DN mm
- AL it still fucking leaks (WSC#DPC066)

Structurally, the constructions given in boldface are similar to both IT-clefts (exemplified in the following section), and at the same time, (BE-)extraposition.⁵ In both cases, the clauses have dummy/expletive *it* in subject position, followed by the copula *be*, some constituent (the 'clefted constituent' in a cleft analysis or the 'remainder predicate' of the main clause, under an extraposed analysis), and finally, a subordinate clause (the 'cleft clause' in a cleft or the demoted subject/object clause in extraposition).

The problem lies in deciding which analysis to adopt for the two constructions. Should they be analysed as clefts or as extraposition? What is more, this question raises several related issues. At a more practical level, it is worth considering which cases have the potential to cause these disambiguation difficulties; while on a more theoretical plane, it is not clear why these constructions should present such similar structures in the first place, making it difficult for us to distinguish between them. Is it just coincidence or is there more to it? More importantly, is it always possible to distinguish between clefts and extraposition? If yes, what criteria can be used to this end? If not, is one construction a special instantiation of the other or is the boundary between them fuzzy rather than discrete? (See Hopper and Thompson's seminal 1980 paper and a recent collection of works edited by Aarts *et al.* (2004) on fuzzy grammar for more details.) If on the other hand, there are cases in which clefts and extraposition

cannot be disentangled, what do these cases look like and how do they differ from prototypical clefts and prototypical extraposition?

Being able to distinguish between π -clefts and extraposition is beneficial because the two constructions play different roles in the discourse; that is, they are used for different purposes. π -clefts are a focusing device, highlighting or contrasting particular bits of information. Extraposition, on the other hand, is connected with the avoidance of having complex subjects at the beginning of the sentence and “serve[s] the two principles of end-focus and end-weight” (Quirk *et al.* 1985: 863). This type of construction is particularly relevant to spoken language, where subjects are almost never complex (not even complex nominal phrases, let alone clausal units). We will see more about the various discourse functions of the two constructions in the following section.

However, their discourse functions apart, clefts and extraposition differ in their syntax. Unlike clefts, extraposition may involve predicates other than the copula *be*, i.e., *It helps to think that my mother will be there waiting for us* and *It surprises me how seriously she takes this stuff*. Furthermore, extraposition allows a greater variety of extraposed elements (NPs, PPs, and so on) whereas clefts only allow a cleft clause in the final position (see Table 1 and example 23).

The questions formulated above form the basis for the work presented here and will be discussed with reference to examples from spontaneous, unplanned face-to-face conversations from the WSC. In spite of the use of spoken data, the results obtained are assumed to apply in equal manner to written language, unless otherwise indicated. The paper has three major aims:

- first, to support the view that π -clefts and extraposition⁶ are distinct and can be distinguished from one another, which is assumed by some (but only stated explicitly by Pérez-Guerra 1998),
- secondly, to draw attention to difficulties in distinguishing the constructions in spoken English, and
- thirdly, to provide an objective test for distinguishing between them, a test which appears to be informally used by some but not explicitly stated in the literature.

The paper is organized as follows. First, extended definitions and examples of π -clefts and extraposition are given. In the following section, further problematic examples are examined, where it is difficult to distinguish between the two constructions. These are used to show that the only recent criteria (proposed by Pérez-Guerra 1998) for differentiating them are not adequate to resolve ambiguities. Pérez-Guerra’s paper gives a diachronic account of the increase in use of

clefted and extraposed constructions throughout the history of the English language as grammaticalized versions of right-dislocation. As part of this account, he proposes a set of criteria for distinguishing clefting and extraposition, which will be outlined in section 3.1. However, it will be argued that these criteria are not sufficient for eliminating ambiguities between the two constructions. A ‘transformational-based test’ (word order re-arrangement test) is proposed as a test for distinguishing between IT-clefts and extraposition, which will be applied to further examples from the WSC, including the earlier constructions given in (A) and (B). It will be shown that the test can be used to successfully distinguish between clefting and extraposition, though there are some cases where stylistic difficulties arise. (The test is, however, based on grammatical acceptability, not stylistic judgments.) Such stylistic problems prove relevant only to cases when both the pre-copula material and the post-copular constituents are short/light. This reinstates earlier questions regarding the discourse function of extraposition. Furthermore, I will also discuss problems related to language medium, in particular difficulties in analysing the structure of speech which arise from the lack of syntactic integratedness of spoken data. Finally, the paper concludes with a summary.

One final remark concerning the theoretical framework used is in order before proceeding with the analysis. The test proposed here has its basis in the (revised) standard models of transformational grammar, but is not conceived of in this work as anything more than a useful tool for the analyst. The current paper is quite neutral with respect to the debate about the merits or demerits of any particular models of generative grammar, transformational or non-transformational.

2 Background

As far as the literature is concerned, cleft clauses and extraposed constructions have been treated separately for the most part, with the exception of a paper by Pérez-Guerra (1998) discussed later in the paper. Furthermore, while there is a wealth of recent research discussing various cleft types (Collins 1987; Hedberg 1988; Delin 1989; Collins 1991; Delin and Oberlander 1995; Oberlander and Delin 1996; Weinert and Miller 1996; Hedberg 2000; Johansson 2001; Lambrecht 2001; Herriman 2003; Collins 2004; Herriman 2004; Delin and Oberlander forthcoming; and others), there is much less current work on extraposition (some of the most recent papers include Pérez-Guerra 1998; Seppänen 1999; and Herriman 2000). Each construction type is considered in turn, starting with clefts.

2.1 *It*-clefts, structure and discourse function

It-clefts are focusing constructions, in which typically a simple sentence (though complex sentences can also be involved) is ‘cleaved’ such that the pronominal *it* appears in initial/subject position, followed by the copula *be*, the clefted constituent which expresses the highlighted or focused element, and finally, the cleft clause, modifying the clefted constituent (see descriptions by Hedberg 1990; Weinert and Miller 1996; and Huddleston and Pullum 2002). Example (1a) gives such a simple sentence, and (1b) illustrates the cleft which can be constructed for the purpose of focusing or highlighting the subject noun phrase *a bright yellow jumper*:

(1a) Henry bought a bright yellow jumper yesterday.

(1b) It was **a bright yellow jumper** that Henry bought yesterday.

The status of the pronoun *it* in cleft constructions has caused some debate in the literature. While some argue that it has referential status (Bolinger 1970; Gundel 1977; Declerck 1988; Hedberg 2000), others contend that it is simply an empty place-holder, devoid of any referring role (Postal and Pullum 1988; Haegeman 1991).

As far as the highlighted element is concerned, there is still disagreement about what types of constituents are allowed to appear in this position. According to Biber *et al.* (1999: 959), the slot can be filled by noun phrases, prepositional phrases, adverbial phrases or adverbial clauses. However, alongside these, Huddleston and Pullum (2002: 1418–1419) add the following possibilities: finite and/or nonfinite clauses (*It's that he's so self-satisfied that I find off-putting* or *It's certainly not to make life easier for us that they are changing the rules*), and adjectival phrases (*It's not lonely he made me feel – it's angry* and *It wasn't green I told you to paint it*).

The exact status of the cleft clause has similarly provoked debate, with opinions ranging from those arguing strongly for its analysis as a relative clause (Hedberg 1990; Huddleston and Pullum 2002) to those still holding notable differences between relative clauses and the nature of cleft clauses⁷ (Quirk and Greenbaum 1985; Miller 1996; Miller and Weinert 1996; Miller 1999; Biber *et al.* 1999). However, it suffices to say that most studies converge on the idea that cleft clauses are at least reminiscent of, even if not identical with, relative clauses.

The discourse function of clefts is to focus new or contrastive information expressed by the clefted constituent (cf. Biber *et al.* 1999 and Lambrecht 2001

among others) and in some cases they can also have a ‘remind me’ role (see Weinert and Miller 1996).

2.2 *Extraposited clauses, structure and discourse function*

We now turn our attention to extraposition. Extraposition is used to ‘lighten’ the load of a sentential subject (and less frequently, an object) by demoting a subordinate clause from subject (or object) position to the end of a sentence. For instance, the subject clause in (2a), *That he left in such a hurry*, is extraposed to the end of the sentence as shown in (2b), with the help of the inserted pronoun *it*. Further examples of extraposition are given in (3)–(5):

- (2a) **That he left in such a hurry** is no surprise.
- (2b) It is no surprise **that he left in such a hurry**.
- (3) It never ceases to surprise and horrify me **how these criminals get away**.
- (4) Well, I believe it to be a crime **to let any murderer walk free like that**.
- (5) The professor found it incredible **that any student could write such an essay on their own**.

There is consensus in the literature regarding the type of clauses which can be extraposed: they can be finite (the most frequent type), or non-finite (with gerunds, as in example (6), much less frequent than infinitives, see (7)), and they can be introduced by *that*, a *wh*-word or nonfinite *to* (Collins 1994; Biber *et al.* 1999; Huddleston and Pullum 2002).

However, one aspect which still causes debate concerns the status of *it*. As with *it*-clefts, *it* has been analyzed in different ways. Some argue that extraposed constructions have two subjects, *it* being the ‘formal’ subject and the extraposed clause being the ‘notional’, ‘real’, or ‘postponed’ one (Jespersen 1972 and Quirk *et al.* 1985); others maintain that *it* is the only subject, the extraposed clause being stripped of its subject-like properties once it is demoted to sentence-final position (Huddleston 1984; Seppänen, Engström and Seppänen 1990; and Seppänen 1999). This issue is beyond the scope of the present paper; what is noteworthy is the similarity between clefting and extraposition regarding the ‘slippery’ nature of *it* in both constructions.

With regard to the discourse function of extraposition, a recent study by Rowley-Jolivet and Carter-Thomas (2005) found two main and partially related uses of extraposition. First, it helps “increase dynamism” by placing new infor-

mation sentence-finally. Secondly, it is used to express the speaker's/writer's evaluative opinion in a "rhetorically effective" way (Rowley-Jolivet and Carter-Thomas 2005: 51; also cf. Hoey 2000; and Hewings and Hewings 2002) by introducing the evaluative comments sentence-initially: *It is obvious that [...]* or *It is unusual that [...]* so that they are less conducive to being challenged. These findings are in agreement with earlier work by Collins (1994) and Herriman (2000).

Despite the fact that extraposition has received most attention in past work, it could be argued that the opposite phenomenon (i.e., non-extraposition, as given in example (2a)) merits equal, if not more, attention. There are two, related reasons for this. First, as shown by Biber *et al.* for *that*-clauses (1999: 676), Collins (1994: 14), Herriman (2000: 584), Mair for *to*-infinitives (1990: 30), and Quirk *et al.* (1985: 1392), from a statistical standpoint, extraposition is more frequent than non-extraposition. Secondly, extraposition is functionally motivated, whether one adopts a psychologically oriented explanation whereby sending heavy constituents to the end of the sentence allows for easier processing (see Langacker 1974: 653 and Huddleston 1984: 354 for this view), or whether it is simply assumed that end-focus and end-weight principles are at work here (see Collins 1994: 15–16 for a more in depth discussion). In other words, both functionally, as well as statistically, extraposition appears to be the norm rather than the exception.

3 *The problem of distinguishing between IT-clefts and extraposition*

The previous section introduced the two constructions of IT-clefts and extraposition. As mentioned earlier, the present paper argues that the two constructions can be reliably distinguished from each other.

It turns out that only one specific type of extraposed clause creates problems in this regard. The term extraposition will be used for the remainder of the paper to refer to this special construction. The label 'extraposition' denotes, in this case, sentences whose extraposed subject clauses involve the lexical verb *be* as the main verb of the superordinate clause, as exemplified in (6) and (7):

- (6) It is pointless **complaining to the head manager.**
- (7) It was very unusual **to see someone so intelligent wasting their time in this way.**

Although extraposition is often contrasted with right dislocation (Huddleston 1984 and McCawley 1988 in Collins 1994: 12–13), with varying degrees of suc-

cess (see Collins 1994: 12–13 for a discussion of why the two constructions cannot always be reliably distinguished from each other), the possible overlap between extraposition and clefting is yet to be addressed, with one exception discussed in what follows.

3.1 Differences between *it*-clefts and extraposition: Pérez-Guerra (1998)

Despite their separate treatment in the literature, *it*-clefts and extraposition exhibit a number of affinities. First, as Pérez-Guerra correctly points out (1998: 8–9), they are both thematically marked, in that the theme slot is occupied by the pronoun *it*. Secondly, the two constructions have similar structural properties: they both have *it* as their initial element, followed by the copula *be*, an additional constituent (the highlighted element in clefts, and the remaining part of the predicate of the main clause in extraposition) and a subordinate clause. Compare the following pair of examples taken from the WSC:

- (8) FE but he does seem to flit around doing little bits and pieces
 MJ well he was over in england <pause> a few years ago <pause>
 and he <pause> went on a special eye course there to qualify
 <pause> **it was very hard for him to get taken** because of the
 old school tie in england and he being from the colonies but er
 after the course he got top marks <pause> so you know
 (WSC#DPC002)
- (9) AW well you've been um going out so much on friday nights you
 wouldn't have taken much notice of that friday night meeting it's
 always auckland anyway auckland always meets on fridays **it's
 a very only the very odd occasions someone else works** you
 know actually <longer pause> someone else actually um
 (WSC#DPC032)

As discussed previously in relation to examples (A) and (B), it is not straightforward to work out whether the two constructions given in (8) and (9) are *it*-clefts or extraposition. In both cases, the clauses involve the pronominal *it*, the copula verb *be* and a sentence-final subordinate clause (*to get taken* and *someone else works*).

A recent paper by Pérez-Guerra (1998: 10–11) claims that the two constructions are distinct. Three criteria are cited as distinguishing between them:

(1) In IT-clefts, the element which follows the verb and precedes the final constituent in the superordinate clause is compulsory, whereas in extraposition, this element is optional.

(2) In the case of *that*-clauses, the subordinate *that*-clause is ‘complete’ in extraposition, but ‘gapped’ in IT-clefts. The gapped element in IT-clefts is coreferential with the obligatory element following the verb and preceding the sentence-final constituent mentioned in (1). Huddleston and Pullum (2002) exemplify this with the cleft construction *It’s the president [I’m referring to __]* (ex. 11i: 1418), which has a gap where the clefted constituent *the president* would normally occur in the unclefted equivalent *I’m referring to the president*. Another example from the WSC Corpus is given in (10), where the clefted constituent *human eyes* functions as an argument of the cleft clause *he was working on* and thus leaves a gap inside it which would be *He was working on human eyes* in the unclefted counterpart. Contrast this with example (11), where the extraposed clause *was so lucky* does not leave such a gap (that is, *I got this* is ‘complete’ without it):

(10) FE i got a postcard from him the other day it looked a BEAUtiful university and um it’s obviously their break over there for about nine weeks and that’s when this course is running and he said there were vets and doctors evidently cos **it was human eyes he was working** <laughs> **on** when he wrote from all over the world there
(WSC#DPC022)

(11) MD unle unle as long as you’ve booked to at i mean i’ve we haven’t i mean we haven’t **it was so lucky i got this** cos the the ballot closed three months ago <pause> just for September
(WSC#DPC023)

(3) According to Pérez-Guerra, and as noted previously by Quirk *et al.* (1985), there are fewer types of clauses which can be clefted than there are clause types which can be extraposed. And indeed, these need not be necessarily clauses, but can also be complex phrases. IT-clefts can have *that*-clauses, *wh*-clauses, and very rarely infinitive clauses, whereas extraposed elements can be: *that*-, *whether*-, *if*-, and *why*-clauses, adverbial clauses, NPs, and PPs. These possibilities are exemplified in Table 1.

Table 1: Examples of the types of clauses which can be clefted and/or extraposed

Clause Type	It-clefts	Extraposition
That-clauses	<i>It was a play that I saw.</i>	<i>It was amazing that he did it.</i>
WH-clauses	<i>It's their arrogance which they hated most of all.</i>	<i>It was amazing how he spoke.</i>
Nonfinite clauses	<i>It's still Mark to come.</i>	<i>It is impossible to please him.</i>
Adverbial clauses	n/a	<i>It seemed hours before she arrived.</i> (ex.10, 11)
Whether-clauses	n/a	no examples given by Pérez-Guerra
If-clauses	n/a	no examples given by Pérez-Guerra
Why-clauses	n/a	no examples given by Pérez-Guerra
NPs	n/a	<i>It is trewe euery word that is wretyn in Brides boke. (Kempe, The Book of Margery Kempe I, 1438, 47, ex. 14, 11⁸)</i>
PPs	n/a	<i>It is better vpon a scaffolde than vpon the grounde. (Fitzherbertm The book of husbandry, 1534, 38, ex. 15, 11)</i>

3.2 Counterexamples from the WSC to Pérez-Guerra (1998)

While Pérez-Guerra (1998) is right in claiming that clefting and extraposition are distinct from each other and that they can be distinguished, his criteria are not particularly successful in doing this, at least as far as some of the data in the WSC are concerned. I consider each of three criteria presented above in turn, and give examples showing their inadequacy in resolving ambiguity between the two constructions.

First, while it is true that many *IT*-clefts exhibit an overt element between the predicate of the superordinate clause and the subordinate clause, these elements need not be present. Some clefts have no such element in that position, as in examples (12) and (13):

(12) WE oh it's so easy i'm sure **it was that well i'll** <laughs> **i'll get my b a in maori tomorrow** <laughter>

IB but the thing is it was the teachers that made us speak er <drawls> er <pause> all say English (WSC#DPC004)

(13) AR well there's auras people have auras and things like that so that er that didn't frighten you either that thing

BT read it mm yeah <unclear word > NO it was just so strange but i didn't TELL the kids cos i didn't want THEM to be frightened and there so **it wasn't like robbie was saying it because he'd heard me talk about it** because i was very careful not to let them hear about it and so once he said it i told him what i'd seen too er so i don't know and i don't know where the ball of energy comes from or anything about it (WSC#DPC121)

The examples above give a particular type of cleft, often termed 'inferential cleft' (see Delahunty 1981; Lambrecht 2001); what is highlighted is not the clefted constituent, but rather the cleft clause, namely *i'll get my b a tomorrow* [*b a* = BA, Bachelor of Arts] in (12), and *robbie was saying it was because he'd heard me talk about it* in (13), respectively. Some (Hedberg 1990; Delahunty and Gatzkiewicz 2000) have analyzed the subordinate clause following the copula as being the clefted constituent itself (rather than the cleft clause). The debate as to whether the subordinate clause should be analyzed as a cleft clause or as a clefted constituent goes beyond the scope of the present work; however, it serves to draw attention to the fact that the analysis of the inferential construction is not a closed case and deserves further investigation.⁹ Even if the subordinate clause is analyzed as a clefted constituent, the construction still remains problematic for the test proposed, because under this analysis, the final constituent coincides with the obligatory constituent (it is not that the element following the verb and preceding the final constituent is obligatory as the test would predict of clefts, but rather, there is no distinct final constituent).

Conversely, in the case of extraposition, the element occurring between the superordinate clause predicate and the subordinate clause is typically not optional, but rather obligatory, as given in (14a) – note that removing it produces ungrammaticality, as indicated in example (14b):

(14a) MK i don't know what it does to the cells of the things it's good for cooking and eating but it dena

CY you're saying **it would be better to put it into that thing and put hot water round the outside** and that way it would be sitting in water

MK microwaves (WSC#DPC077)

(14b) * It would be to put it into that thing and put hot water round the outside.

As far as the second criterion is concerned, clearly it is not only *that*-clauses that can potentially cause ambiguity problems; but other clauses (such as non-finite, *WH*-clauses, etc., see examples (B), (8) and (9)) can also be involved in problematic cases. However, leaving this point aside, it is not always the case that clefted constituents function as arguments of the cleft clause which they relate to. That is, in some cases, the cleft clause is not exactly a relative clause and the clefted constituent is rather an adjunct of the cleft clause. In such cases, the gap test does not hold in the same way as we have seen earlier, since the cleft clause is ‘complete’ without the ‘missing’ clefted constituent. Consider example (15a):

- (15a) TS and i think i'll be i'm sure i'll get maturity onset diabetes
 KA **it's for the sugar that it has to** <pause> **secrete the insulin**
 <latch>
 TS yeah but i think your pancreas isn't that what they th that's one
 of the theories why they think people get maturity onset diabetes
 just from years of <pause> pancreas getting tired
 (WSC#DPC024)

The clefted constituent *for the sugar* is an adjunct of the cleft clause *it has to secrete the insulin* since it is optional and the clause is complete without it (admittedly, the cleft clause does allow the PP to be present, but it does not require it). Hence in examples such as (15a), the gap test may be considered a weaker and perhaps not entirely convincing means for establishing the desired cleft classification. Similar examples can be found of cases where the clefted constituent functions only as an adjunct, and not as an argument in the cleft clause, and thus leaves a ‘weaker’ gap in it. Three additional ones are given in (15b)–(15d) from the WSC:

- (15b)BT oh <laughs> oh
 AL and just at that point you suddenly get the shits you know you
 think i'm in a car with a mad man <laughs>
 BT <laughs> that like <hums scary movie music> de da de da de da
 <laughs>
 AL that's right <laughs> but you s **it's just at that point you realise**
how out of how little control you've got as a passenger in a
car <pause> and um so then i'm starting to think about what to
 do next because what's going to happen is ac is a another curve
 coming up <latch>
 (WSC#DPC049)

- (15c) LR and do you all have a um <pause> a support group time
 QT my chest yes <pause> yep
 LR at work
 QT yep it's all work time
 LR is that part of their programme
 QT that's part of the programme and **it's in six weeks we'll get our first certificate**
 LR <with high pitched voice> <drawls> mm <latch>
 QT we'll all get a little certificate and and at six months we get another certificate and then at the end (WSC#DPC334)
- (15d) XX um his girlfriend had shouted him a night at the plaza apparently they had a special deal on or it might have been a weekend even and so there they were at the plaza hotel all flash and everything and he thought oh who can i ring <laughs> so he rang his mum and she said **it wasn't until afterwards she thought or one of her friends said** and he rang YOU and she thought oh yeah </laughs> <laughs> it might be a bit of a strange thing to do ring your mum <latch> (WSC#DPC334)

Admittedly, cases such as the ones exemplified in (15a)–(15d), are infrequent in the data (particularly in spoken data); however, they can still cause ambiguity problems.

Finally, Pérez-Guerra makes the interesting remark that clefts involve a more limited distribution of clause types, in comparison with extraposition, which can exhibit a much larger variety of clause types (see Table 1). Unfortunately, however, this observation does not help distinguish the constructions since the majority of the clause types which can only be found in one construction (extraposition) are themselves very rare. In other words, while extraposition can involve extraposed adverbial clauses, *whether*-clauses, *if*-clauses, *why*-clauses, NPs and PPs, as Pérez-Guerra himself concedes (1998: 12), these types of extraposition are themselves rarely used. For instance, in the LOB corpus data he investigated, only 1.3 per cent of clauses are final adverbial clauses, even less are NPs (1.1%), and hardly any are PPs (0.5%). This means that while the remark is worth noting, it is in itself not widely applicable as a test in clarifying potential ambiguity between *it*-clefts and extraposition.

It appears that despite making interesting predictions regarding the nature of the two constructions in general, and trends which each one might exhibit (such as optional versus obligatory elements after the superordinate clause predicate, gapping versus complete clauses, and various clause types which can be found with one construction but not the other), Pérez-Guerra (1998) does not offer a solution to our earlier problem exemplified in (A) and (B), at the beginning of this paper. In the next section, a transformational-based test is proposed as a more widely applicable and objective test in distinguishing between clefting and extraposition.

4 Resolving ambiguity between *it*-clefts and extraposition

4.1 The test: A transformational-based test

The structural similarity in the two constructions can be represented symbolically in the following way:

***it*-cleft**

It + *be* copula + clefted constituent + cleft clause

extraposition

It + *be* copula + remainder predicate + extraposed clause



 = superordinate clause predicate

The key to the test proposed here for distinguishing the two constructions lies in the crucial observation concerning the status of the copula. If we take a basic declarative sentence and convert it to a cleft or an extraposed construction, the status of the copula is different in the two newly obtained sentences. In the case of extraposition, the copula *be* is actually part of the superordinate clause predicate, whereas in the cleft construction, the copula is inserted (just as the pronominal *it* is inserted), in addition to the original sentential predicate.

It can, then, be predicted that the process of *un-cleaving* a cleft construction by eliminating the inserted subject *it* and shifting the clefted constituent to sentence-initial position (while leaving the copula *be* and the cleft clause as they are) will result in ungrammaticality. In contrast, *reinstating* an extraposed clause to its original subject (or object) position will not produce such ungrammaticality since the copula was part of the original matrix clause predicate in the first place. Hence, the test for distinguishing between clefts and extraposition is outlined below:

Test for distinguishing between clefts and extraposition

- 1) Eliminate *it*.
- 2) Move sentence-final clause to the front of the sentence.
- 3) If the construction obtained is grammatical, then the original structure is an example of extraposition. If, on the other hand, the construction obtained is ungrammatical, then the original structure is a cleft.

Examples (16) and (17) show that the test classifies correctly the various clefts and extraposition structures discussed in earlier sections:

(16) Applying the transformational-based test to IT-clefts

cleft It was a bright yellow jumper that Henry bought yesterday.

un-cleaving * That Henry bought yesterday was a bright yellow jumper.

cleft It's that he's so self-satisfied *that I find so off-putting*.

un-cleaving * That I find so off-putting is that he's so self-satisfied.

cleft It's not lonely *he made me feel*. (It's angry.)

un-cleaving * He made me feel is not lonely.

cleft It wasn't green *I told you to paint it*.

un-cleaving * I told you to paint it wasn't green.

(17) Applying the transformational-based test to extraposition

extraposition It is no surprise *that he left in such a hurry*.

reinstatement That he left in such a hurry is no surprise.

extraposition It never ceases to surprise me *how these criminals get away*.

reinstatement How these criminals get away never ceases to surprise me.

extraposition I believe it to be a crime *to let any pedophile walk free like that*.

reinstatement I believe (that) to let any pedophile walk free like that is a crime.

extraposition It is pointless *complaining to the head manager*.
reinstatement Complaining to the head manager is pointless.

So what about the earlier problematic examples? Their analysis is given below:

Applying the test to (A):

... it's quite good *to hear that faith's actually experimenting with a few different things ...*

To hear that faith's experimenting with a few different things is quite good.

Grammatical¹⁰ → result: EXTRAPOSITION

Applying the test to (B):

... it's about six weeks ago *they did the job*

* They did the job is about six weeks ago.

Ungrammatical → result: CLEFT

Applying the test to (8):

... it was very hard *for him to get taken ...*

For him to get taken was very hard.

Grammatical → result: EXTRAPOSITION

Applying the test to (9):

... it's only the very odd occasions *someone else works ...*

* Someone else works is only the very odd occasions.

Ungrammatical → result: CLEFT

4.2 *Other issues and problematic cases*

Despite the general applicability of the test, it is not always possible to reconcile all problematic examples. Sometimes, this is because of the very nature of spoken language,¹¹ which is syntactically loosely integrated, see example (18):

(18) FR i'd better try swallowing a piece when i'm before it's cooked
 since that's what the chemist reckons that's what that's what
 estroscopy's like <laughs>

MK HANG on

DV <laughs>

MK why don't we stick that down have a look down there

DV oh yeah it's got a hole

MK <laughs> mm

DV i wonder if you can see right <laughs> probably be perfectly strange

MK you can't even blow through it

DV mm you can't either <sniffs> <pause> i thought it had a hole right through

MK **it's only the beginning and the end the worm gets lost** <laughs> it's still in there <longer pause>

DV oh well <longer pause> <clears throat> <longer pause> well some of us have got things to do <longer pause>

(WSC#DPC014)

It is not clear whether the construction *it's only the beginning and the end the worm gets lost* contains two disconnected clauses, the first, *it's only the beginning and the end*, being a response to the previous speaker's utterance, and the second a new, unrelated clause, or whether the construction is one long 'sentence' containing two, un-integrated clauses. If the latter analysis is adopted, the test gives **The worm gets lost is only the beginning and the end*, which is clearly ungrammatical, and leads to the conclusion that we are dealing with a cleft. However, if the original construction were more tightly integrated syntactically (as one might find for instance in a written text), i.e., *It's only in the beginning and in the end that the worm gets lost*, then the test would give *?That the worm gets lost is only in the beginning and in the end*, whose grammaticality is not so easily assessable. The example illustrates a broader difficulty, namely how to analyse spoken language, and the potential problems which stem from the assumption that spoken data has a relatively high degree of syntactic tightness, as found in writing (see Chafe and Danielewicz 1987 for further details).

A different type of syntactic looseness is found in examples where the extraposed clause is not integrated inside the clause complex which it is part of, as in (19):

- (19) AL and then when he's that was pretty shattering and then after i'd actually two years later after i'd had them all fixed up and everything he took other moulds and **to actually compare them it was just <pause> unbelievable** (WSC#DPC062)

Here, the extraposed clause *to actually compare them* is topicalized, preceding the clause *it was just unbelievable*, which is in fact the opposite of what extrapo-

sition is actually used for. Note that the presence of *it* indicates that we are dealing with some form of extraposition – but the function of construction in discourse is not the usual one in this case.

A different analysis would be to take the anaphoric *it* as pointing further back to the clause *he took other moulds* (it could not point only to the noun phrase *other moulds*, since the predicate is *was*, and the noun phrase is plural) rather than to the non-finite clause *to actually compare them*. This way, the clause *it was just unbelievable* is independent from the preceding clause, and thus we do not have a case of extraposition. It is simply unclear what clause is co-indexed with anaphoric pronoun.

It is not only extraposed clauses that can appear syntactically loosely integrated, or rather un-integrated, in spoken language. Cleft constructions also suffer from the same phenomenon. Consider example (20):

- (20) BA what was that one that you made
 TR yeah **it was this little thing just this folding that they do** and they wrap <pause> this little thing around them and they put a strip of dried meat in it apparently according to the book anyway
 (WSC#DPC025)

The relative clause *that they do* modifies *this little thing just this folding*, which is in fact not one noun phrase but rather two noun phrases, treated as one single one and produced by the speaker online, while she was still thinking about what she wanted to say.

Similarly, the speaker in (21) starts to form what looks like an *it*-cleft, but whose relative clause is not introduced by the expected relativizer *that*, but appears instead as a temporal adverbial clause introduced by *then*. This is presumably because of the complexity of the construction, where the clefted constituent is made up of two coordinated clauses:

- (21) LL **it's only when HE says that it's important or when he's seen it and then HE talks about it then it's important** but if you talk about something that he doesn't know anything about then it's not important
 (WSC#DPC007)

There is another possible analysis of the example in (21); namely the *when*-clause could be taken to be a headless relative and the *it ... then* to form a correlative construction. This serves to illustrate again the general problems encountered when dealing with spoken data, that is, the availability of different analyses and the lack of a clear-cut solution.

Finally, a common feature of spoken language is anacoluthon, where a phrase is used in two, disconnected clauses.¹² For instance, in (22), the noun phrase *mother* is used in the cleft as a highlighted element (the clefted constituent), while at the same time functioning as subject in the following clause *went back to her doctor* (which would constitute the clefted clause in a tightly syntactically integrated cleft *It was my mother **that** went back to her doctor and ...*):

- (22) AN see m mum and dad had hepatitis <pause> and they go to different doctors and one of them gave <pause> just one sort of mal malarial tab <pause> tablet and the other one gave two sorts of malarial tablet so then <clears throat> i think **it was mum went back to her doctor and asked for the** <laughs> **other sort too**
(WSC#DPC008)

False starts or restarts can also permeate spoken language as in (23):

- (23) KA yeah straight sorry <pause> straight feminists
TS well i know especially her because she's ULTRA political <latch>
LU well
KA yeah but <pause> **it's sort of weird the way straight women the way they put down men so much** i suppose lesbians don't talk about them so much <latch>
TS no cause we don't have to even worry about the politics of men <pause>
(WSC#DPC024)

The sentence *The way straight women the way they put down men so much is sort of weird* does not seem to be grammatically 'acceptable' in writing, but it is in speech.

This point brings us to a final observation concerning the distinction between grammatical acceptability and stylistic preference. In some cases, extraposed clauses are not as stylistically 'accommodating' when being reinstated to subject position. The transformation produces constructions which are stylistically awkward, though still grammatically acceptable. However, in spite of this stylistic awkwardness, the test still gives the desired results when applied in these instances. For example, in (24), the reinstated extraposed clause *having a fight is worth(while) i suppose* may not be stylistically preferred, but it is nevertheless grammatical:

(24) LN yeah it was the fact that he was going to lose in the long yeah he's gonna lose in the long run anyway cos of course they will be allowed to get away with it <laughs> **it's worth having a fight** i suppose whereas like with mum we pay you know well over half

MR how much do you pay

LN seventy (WSC#DPC090)

This issue seems to affect constructions where both the predicate left over from the superordinate clause and the extraposed clause are particularly short/light. For instance in (25) and (26), the predicate 'remainders' *good* and *very hard*, respectively, and the extraposed clauses *that you do* and *to fit* are all short/light:

(25) BB you're not jen

JN no <laughs> i'm just a dick < laughs>

SS <laughs> no no no you you **it's good that you do**

JN oh

BB is that why dave likes you (WSC#DPC093)

(26) MR we'd load the boot up we had a vanguard with a curved lid boot lid **it was very hard to fit** <latch>

XX what's a vanga oh the guard (WSC#DPC063)

This is not so much of a problem for extraposition where either only the remainder of the predicate is short/light, as in (27), or in cases where only the extraposed clause is short/light, as in (28). In both examples, the reinstated extraposed clauses are stylistically acceptable: *that there's different kinds of fits is interesting* in (27) and *to do it was jolly decent of her* in (28):

(27) JM no

XX but **it's interesting that there's different kinds of fits** <long pause>

CR yeah people can some people grow out of it too some people you know who fit as children don't at all (WSC#DPC070)

- (28) LC it was quite <laughs> funny having having all those photos taken with this sneering and smiling sarah behind the photo behind the camera <laughs>
- MG what <laughs>
- LC **it was jolly decent of her to do it** because she's actually very busy at the moment <longer pause> and she's being quite devoted to it (WSC#DPC039)

Instead, it seems that stylistic problems appear only when both the extraposed clause and the leftover predicate are short/light. This is somewhat surprising in light of the various explanations put forward in the literature regarding the motivation for extraposition. If the use of extraposition is indeed connected to processing ease, as suggested by Langacker (1974) and Huddleston (1984), and then a light subject clause should be perfectly acceptable, both grammatically and stylistically. So why are short/light clauses extraposed in the first place? Moreover, why should the reinstatement of the extraposed clause in its original position be stylistically awkward in such cases? All these examples raise further questions about the nature and properties of the constructions discussed and indicate that there is still much work to be done towards a better understanding of the discourse function of extraposition.

5 Conclusion

IT-clefts and extraposition share various functional and structural properties. They are both thematically marked, clefting places a given constituent in focus position, and extraposition moves a clause out of subject (theme) position and places *it* sentence-finally. Similarly, both constructions have pronominal *it* in initial subject position followed by the copula *be* and a sentence-final clause (cleft clause in clefts and extraposed clause in extraposition). These affinities raise questions as to the basis/criteria on which previous studies have each chosen the cleft and extraposed examples analyzed, respectively, and whether they had 'masked' difficulties in deciding what to count as a cleft or an extraposed clause.

Despite their affinities, however, clefting and extraposition can be distinguished from each other in both, spoken and written language. The test proposed here for differentiating them involves a word order re-arrangement test, which states that clefts result in ungrammaticality when the cleft clause is pre-posed to

replace *it*, while extraposition produces grammatical structures when the extraposed clause is reinstated in subject position.

Clefting and extraposition play a key role in the organization of discourse and the packaging of information, and give insight into the way speakers/writers construct linguistic expression. It is therefore of practical, as well as theoretical importance, to recognize similarities and differences between the two constructions. While this work provides some answers, problems still remain, such as stylistic issues, difficulties regarding language medium, and questions relating to the specific function of extraposition in discourse.

Notes

1. Author's note: This paper would not be possible without Jim Miller, who took part in long discussions of how to distinguish clefts and extraposition, and read drafts and revisions of the paper. Similarly, I am indebted to Frantisek Lichtenberk for his meticulous draft reading which he so patiently endured. I am also grateful for the comments made by the referee which also helped improve the paper. Any remaining mistakes are, of course, my own.
2. More information on the corpus can be found at: <http://www.vuw.ac.nz/lals/corpora/index.aspx#wsc>.
3. The examples provided here are mostly from the Wellington Spoken Corpus of New Zealand English, as indicated by the file names given directly after each example (i.e., WSC#DPC002). All others are my own (where no such references are given).
4. The WSC corpus is annotated for various discourse features. The relevant ones (some have been excluded here for clarity and simplicity) have been included in the examples cited, given in angle brackets. A <pause> is less than one second long, a <longer pause> signals a pause between one and two seconds, and anything longer than five seconds is coded by <long pause>. The remaining discourse features coded are self-explanatory (i.e., <laughs> indicates that the speaker is laughing, <coughs> indicates that the speaker is coughing and so on).
5. The only type of extraposition we are dealing with in this paper is *be*-extraposition, as the possible ambiguity between clefting and extraposition only occurs in the case of extraposition involving the verb *be*.
6. The choice of terminology is explained at the beginning of Section 3.
7. The debate regarding the exact status of the relative clause is by no means resolved for the case of written language, but it is even more complex in

spontaneous, spoken language. Examples from Quirk *et al.* (1985: 953) such as *It was because he was ill (that) we decided to return*, and *It was in September (that) I first noticed it* show clauses (*that we decided to return*, and *that I first noticed it*) which do not modify noun phrases, but rather entire clauses (*because he was ill*, and *in September*, respectively). Their argument is based on the lack of a noun as antecedent, an argument which Miller (1999) has argued against. As pointed out in Miller, “the lack of a noun antecedent does not automatically disqualify a sequence as a relative clause” (1999: 17), however, the impossibility of replacing *that* with a WH-word does. Furthermore, in speech, new constructions (not mentioned in Biber *et al.* 1999) such as *It was in September **when I first noticed it*** or *It was in Edinburgh **where we found the picture***, were noted (Miller 1999: 19, examples 31 and 32) which could not be analysed as relative clauses, but rather as headless relatives (see Miller 1999 for a discussion).

8. The examples given by Pérez-Guerra (1998) for extraposed NPs and extraposed PPs are all from Middle English, none being from Modern English. This raises the question of whether these constructions are prevalent, if at all still existent, in English today.
9. Furthermore, it must be said that the inferential construction is more controversial still, in that, although it has been mostly analyzed as a cleft construction (as mentioned earlier), this analysis is not without its problems. Unlike IT-clefts, the inferential has by definition no element following the subordinate clause. This is often explained by ellipsis, as in the following example:

(i) *It is not that he is reluctant to help [that bothers me].*

However, more problematic are examples which go further away from the IT-cleft prototype in that they do *not allow* a cleft clause following the clefted constituent at all as in (ii).

(ii) LL: *oh it's not that i don't want to have to look over my shoulder*
(WSC DPC007)

DV: *oh it's just it just tastes the same it's just that it makes the spaghetti bigg*
(WSC DPC014)

BD: *it's just you know how i had those shin splints it's just that i don't get those <longer pause>*
(WSC DPC037)

The inferential construction also has no focal stress, and a limited number of elements which occur after the verb (*just, like, simply* and *rather*). Another characteristic of the construction is its appearance with the negative particle

not. The use of the negative changes the presupposition inferred by the subordinate clause. Let us reconsider example (i) without the cleft clause, and the positive counterpart of the construction:

(iii) *It is not that he is reluctant to help.*

(iv) *It is that he is reluctant to help.*

The presupposition in (iii) is that he is NOT reluctant to help, whereas in (iv), the presupposition is that he IS reluctant to help. However, adding a cleft clause changes the presupposition in the negative inferential, but not in the positive inferential:

(v) *It is not that he is reluctant to help that bothers me.* PP → He IS reluctant to help.

(vi) *It is that he is reluctant to help that bothers me.* PP → He is reluctant to help.

What all this shows is that inferential construction requires more data and further analysis.

10. It is however worth noting that the occurrence of this type of construction, i.e., *To hear that faith's experimenting with a few different things is quite good*, is not attested in corpora of spontaneous spoken English. In fact, the use of *that*-clauses and *to*-clauses pre-verbally is extremely rare in spoken language, and extraposed *that*-clauses and *to*-clauses are much more preferred, according to Biber *et al.* (1999: 676 and 724, respectively). Furthermore, as noted by Quirk *et al.* (1985: 964–965) some extraposition examples do not allow reverting to the 'un-extraposed' constructions, in either writing or speech, for example extraposed clauses of the type: *It seems / appears / happened / chanced*, as in *It seems that everything is fine* → * *That everything is fine seems*, or *It happened that the old man was walking past again* → * *That the old man was walking past again happened* (ibid).
11. Note that this would not constitute a problem for written language.
12. The term is well known in the community, but its origin is somewhat unclear and I was not able to find any references in connection with it.

References

- Aarts, Bas, David Denison, Evelien Keizer and Gergana Popova (eds.). 2004. *Fuzzy grammar – a reader*. Oxford: Oxford University Press.
- Biber, Douglas, Stig Johansson, Geoffrey Leech, Susan Conrad and Edward Finegan. 1999. *Longman grammar of spoken and written English*. Harlow: Pearson Education.
- Bolinger, Dwight. 1970. The lexical value of *it*. *Working Papers in Linguistics* (University of Hawaii) 2: 57–76.
- Chafe, Wallace and Jane Danielewicz. 1987. Properties of spoken and written language. In R. Horowitz and S. J. Samuels (eds.). *Comprehending oral and written language*, 83–113. San Diego: Academic Press.
- Collins, Peter. 1987. Clefts and pseudo-cleft constructions in English spoken and written discourse. *ICAME Journal* 11: 5–17.
- Collins, Peter. 1991. *Cleft and pseudo-cleft constructions in English*. London: Routledge.
- Collins, Peter. 1995. Extraposition in English. *Functions of Language* 1: 7–24.
- Collins, Peter. 2004. Reversed *what*-clefts in English: Information structure and discourse function. *Australian Review of Applied Linguistics* 27: 63–74.
- Declerck, Renaat. 1988. *Studies on copular sentences, clefts and pseudo-clefts*. Leuven: Foris.
- Delahunty, Gerald. 1981. Topics in the syntax and semantics of English cleft sentences. PhD dissertation, University of California.
- Delahunty, Gerald and Laura Gatzkiewicz. 2000. On the Spanish inferential construction *ser que*. *Pragmatics* 10: 301–272.
- Delin, Judy. 1989. Cleft constructions in discourse. PhD dissertation, University of Edinburgh.
- Delin, Judy and Jon Oberlander. 1995. Syntactic constraints on discourse structure: The case of *it*-clefts. *Linguistics* 33: 456–500.
- Delin, Judy and Jon Oberlander. Forthcoming. *Cleft constructions in context: Some suggestions for research methodology*.
See <http://www.gem.stir.ac.uk/delin-publications.html#cleft>; accessed in February 2008.
- Gundel, Jeanette. 1977. Where do cleft sentences come from? *Language* 53: 542–559.
- Haegeman, Liliane. 1991 [1994]. *Introduction to government and binding theory*. Oxford: Blackwell.

- Hedberg, Nancy. 1990. *Discourse pragmatics and cleft sentences in English*. PhD dissertation, University of Minnesota.
- Hedberg, Nancy. 1998. The discourse function of cleft sentences in Spoken English. Paper presented at the Linguistic Society of America Meeting. New Orleans, Louisiana.
- Hedberg, Nancy. 2000. The referential status of clefts. *Language* 76: 891–920.
- Herriman, Jennifer. 2000. Extraposition in English: A study of the interaction between the matrix predicate and the type of extraposed clause. *English Studies* 6: 582–599.
- Herriman, Jennifer. 2003. Negotiating identity: The interpersonal functions of *wh*-clefts in English. *Functions of Language* 10: 1–30.
- Herriman, Jennifer. 2004. Identifying relations: The semantic functions of *wh*-clefts in English. *Text* 24: 447–469.
- Hewings, Martin and Ann Hewings. 2002. *It is interesting to note that ...* : A comparative study of anticipatory ‘it’ in student and published writing. *English for Specific Purposes* 21: 367–383.
- Hoey, Michael. 2000. Persuasive rhetoric in linguistics: A stylistic study of some features of the language of Noam Chomsky. In S. Hunston and G. Thompson (eds.). *Evaluation in text. Authorial stance and the construction of discourse*, 28–37. Oxford: Oxford University Press.
- Hopper, Paul and Sandra Thompson. 1980. Transitivity in grammar and discourse. *Language* 56: 251–299.
- Huddleston, Rodney. 1984. *Introduction to the grammar of English*. Cambridge: Cambridge University Press.
- Huddleston, Rodney and Geoffrey Pullum. 2002. *The Cambridge grammar of the English language*. New York: Cambridge University Press.
- Jespersen, Otto. 1927 [1940]. *A modern English grammar on historical principles*. Copenhagen: Ejnar Munksgaard.
- Johansson, Mats. 2001. Clefts in contrast: A contrastive study of *it*-clefts and *wh*-clefts in English and Swedish texts and translations. *Linguistics* 39: 547–582.
- Lambrech, Knud. 2001. A framework for the analysis of cleft constructions. *Linguistics* 39: 463–516.
- Langacker, Ronald. 1974. Movement rules in a functional perspective. *Language* 50: 630–664.

- McCawley, James D. 1988. *The syntactic phenomena of English*. Volume II. Chicago: The University of Chicago Press.
- Mair, Christian. 1990. *Infinitival complement clauses in English*. Cambridge: Cambridge University Press.
- Miller, Jim. 1996. Clefts, particles and word order in languages of Europe. *Language Sciences* 18: 111–125.
- Miller, Jim. 1999. Magnasyntax and syntactic analysis. *Revue française de linguistique appliquée* IV: 7–20.
- Oberlander, Jon and Judy Delin. 1996. The function and interpretation of reverse *wh*-clefts in spoken discourse. *Language and Speech* 39: 185–227.
- Pérez-Guerra, Javier. 1998. Integrating right-dislocated constituents: A study on cleaving and extraposition in the recent history of the English language. *Folia Linguistica Historica* XIX: 7–25.
- Postal, Paul and Geoffrey Pullum. 1988. Expletive noun phrases in subcategorized positions. *Linguistic Inquiry* 19: 635–667.
- Quirk, Randolph, Sidney Greenbaum, Geoffrey Leech and Jan Svartvik. 1985. *A comprehensive grammar of the English language*. New York and London: Longman.
- Rowley-Jolivet, Elizabeth and Shirley Carter-Thomas. 2005. Genre awareness and rhetorical appropriacy: Manipulation of information structure by NS and NNS scientists in the international conference setting. *English for Specific Purposes* 24: 41–64.
- Seppänen, Aimo. 1999. Extraposition in English revisited. *Neuphilologische Mitteilungen* 100: 51–66.
- Seppänen, Aimo, Clas Göran Engström and Ruth Seppänen. 1990. On the so-called anticipatory *It*. *Zeitschrift für Phonetik, Sprachwissenschaft und Kommunikationsforschung* 43: 748–776.
- Weinert, Regina and Jim Miller. 1996. Cleft constructions in spoken language. *Journal of Pragmatics* 25: 173–206.

