

University of Groningen

Topics in Corpus-Based Dutch Syntax

Beek, Leonoor Johanneke van der

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version

Publisher's PDF, also known as Version of record

Publication date:

2005

[Link to publication in University of Groningen/UMCG research database](#)

Citation for published version (APA):

Beek, L. J. V. D. (2005). *Topics in Corpus-Based Dutch Syntax*. s.n.

Copyright

Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

The publication may also be distributed here under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license. More information can be found on the University of Groningen website: <https://www.rug.nl/library/open-access/self-archiving-pure/taverne-amendment>.

Take-down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): <http://www.rug.nl/research/portal>. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.

Chapter 2

Clefts

This chapter is concerned with the *it*-cleft construction in Dutch. We claim that we have to distinguish two types of cleft sentences: the transitive cleft, which focuses a nominal phrase and contains a relative clause, and the intransitive cleft, which may focus a range of different phrases and contains a final complementizer clause. Formal analyses of both types are presented. An earlier version of this chapter appeared as Van der Beek (2003).

2.1 Introduction

Dutch *it*-clefts are a puzzling construction. They consist of the same basic elements as English clefts—the pronoun *het* ‘it’, the verb *zijn* ‘to be’, the focused phrase (c-focus) and a final clause—but agreement is different: if the c-focus is plural, then the copula is plural too, even though the subject is *het* ‘it’ (1-a)¹. This appears to be in conflict with the otherwise strict subject verb agreement in Dutch.

Accounting for the agreement in Dutch clefts is further complicated by the fact that the argument structure of clefts depends on whether or not the c-focus is a pronoun: if the c-focus is a full noun phrase, *het* is in the canonical preverbal subject position and the c-focus in the canonical postverbal object position (1-a)-(1-b). But if the c-focus is pronominal, then it is in subject position and *het* is in object position (1-c), generally.

- (1) a. Het zijn niet de vliegtuigen die mij uit de slaap houden.
it are not the airplanes that me out the sleep keep

¹Unless explicitly stated otherwise, all (grammatical) examples are from corpus data. We used several newspaper corpora, *Volkskrant*, *NRC Handelsblad* and *Algemeen Dagblad*, and the Corpus of Spoken Dutch (CGN) to retrieve the examples. For rare constructions, we used the web as an additional source of data.

- It's not the airplanes that keep me awake.*
- b. Het is immers niet de trainer die kansen voor open
 it is after all not the trainer who chances in front of open
 doel verknalt.
 goal loses
After all, it's not the coach who misses the easy shots.
- c. Ik ben het die dom doet.
 I am it that stupid does
I am the one acting stupid.
- d. Het was op zijn aandringen, dat ik de redactie van de
 it was on his insistence that I the wording of the
 adviesaanvraag [...] zo heb veranderd.
 advice appeal have thus changed
*It was on his insistence that I changed the wording of the appeal
 for advice [...].*
- e. Het is omdat ik dit voorheb, dat ik hem begrijp.
 it is because I this have-before, that I him understand
It is because I have this advantage, that I understand him.

A second challenge for the analysis of Dutch *it*-clefts is the difference between clefts with final relative clauses (1-a)-(1-c) on the one hand and *it*-clefts with final complementizer clauses (1-d)-(1-e) on the other hand. While the final clause in the first three examples is introduced by the plural/common relativizer *die* (1-a)-(1-c), the clause is headed by the complementizer *dat* in the examples (1-d)-(1-e).² A similar contrast can be observed in English, where the *that*-clause can often be analyzed as a relative clause (in which *that* may be replaced by *who* or *which*), but not always. See also Quirk et al. (1985) for the characteristics of the final clause in English cleft sentences.

All examples in (1) contain a subject *het*, a form of *zijn* 'to be', a focused phrase and a clause, and in each sentence, given information is extraposed in order to focus a certain constituent. Despite these similarities, this chapter argues that the Dutch *it*-cleft is in fact two constructions: one with the transitive (specificational) copula and a discontinuous TOPIC for clefting nominals and one with the intransitive (existential) verb *zijn* 'to be' for clefting other syntactic categories (as well as some nominals). The first has a final relative clause, but cannot be reduced to any other relative clause construction. The second has a complementizer clause, but cannot be reduced to any other

²The complementizer is homonymous to the neuter relativizer. We know that it is in fact the complementizer because there are no neuter nominal 'traces' in the embedded clause.

complementizer clause construction. We account for the agreement patterns in (1) without violating the generally assumed canonical word order rules for Dutch nor subject-verb agreement. In addition, we show how both the argument structures in (1-b) and in (1-c) can be generated by one set of rules.

Section 2.2 presents an analysis of transitive clefts and in section 2.3 it is argued that the intransitive cleft is a separate construction, for which a separate formal analysis is presented. The chapter concludes with a summary and discussion of some open ends in section 2.4.

2.2 Transitive Clefts

The first type of cleft has a final relative clause and a nominal c-focus, which is either a pronoun or a full NP (1-a)-(1-c). The construction has various interesting features: it appears to violate the otherwise strict subject verb agreement, the relative clause appears not to agree with its antecedent if this antecedent is a pronoun and the argument structure depends on the syntactic category of the c-focus.

We shall show that the final clause cannot be reduced to a regular post-nominal or extraposed relative clause modifier. Instead, it must be analyzed as a specific construction for focusing nominals. After investigating the syntactic properties of the c-focus, the subject pronoun and the relative clause, a formal analysis of the construction is presented which accounts for the agreement features without violating the general word order principles of Dutch or the principle of subject-verb agreement.

2.2.1 Differences between cleft clauses and other relative clauses

The relative clause cleft is very similar to predicative copular constructions in which the NP predicate has a postnominal relative clause modifier. This may even lead to ambiguities between the two readings. Compare the two text fragments in (2-a) and (2-b). Both examples contain an almost exact repetition of example (1-b). In (2-a), it is presented in the context of (1-b) in the corpus. This is an example of a cleft construction. It negates the identification of the person who misses the easy shots with the coach, while putting heavy focus on *trainer*. No other part of the sentence is or can be focused. The c-focus and relative clause do not form a syntactic or semantic unit.

- (2) a. [Coach Vonk] weigert de verantwoordelijkheid voor de
 coach Vonk refuses the responsibility for the
 malaise op zich te nemen. Het is immers niet de trainer die
 slump on self to take it is after all not the coach who
 kansen voor open doel verknalt.
 chances in front of open goal misses
*Coach Vonk refuses to take the responsibility for the slump. After
 all, it's not the coach who misses the easy shots.*
- b. Je zal verbaasd zijn te horen wie er ontslagen is. Het is
 you will surprised be to hear who there fired is it is
 niet de trainer die kansen voor open doel verknalt.
 not the coach who chances in front of open goal misses
*You'll be surprised to hear who got fired. It is not the trainer who
 misses the easy shots.*

In the second fragment, the same sentence is placed in a different context. The prosody of the sentence changes: the main stress shifts from *trainer* to *niet*. Furthermore, the meaning is completely different: the (negated) identification is not between the person who misses the easy shots and the coach, but between some third person (the one who got fired) and the coach, of which we know he missed some easy shots. And in contrast to fragment (2-a), the information structure in this sentence is not fixed: though less likely, the focus (and therefore the main stress) could also be on *open*, *goal* or *misses*. Finally, the two differ with respect to their syntax: the relative clause and the predicate nominal in this sentence form a semantic and a syntactic unit (an NP). This second fragment is not a cleft sentence.

The string may contain some clues as to whether an example sentence is a regular post-nominal modifier or a cleft. For example, proper names and pronouns seldom have relative clause modifiers, but they do occur frequently in clefts. Disambiguation is also possible on the basis of constituent structure: if the complement and the relative clause may be topicalized together, as in (3), without a change of meaning, then the two form one NP and the original sentence is not a cleft construction. But sometimes disambiguation is just a matter of interpretation. The corpus examples in this chapter were only included if their contexts showed them to be clear examples of the *it*-cleft construction, not post-nominal modifiers.

- (3) De trainer die kansen voor open doel verknalt is het niet
 the coach who misses in front of open goal misses is it not
It's not the trainer who misses the easy shots.

For differences between the *it*-cleft and other constructions containing an integrated relative clause in English, see Huddleston and Pullum (2002, p.1416). There is an extensive literature on the semantic and pragmatic characteristics that are specific to the cleft construction. Some pointers are Chomsky (1972); Prince (1978); Delahunty (1981); Atlas and Levinson (1981); Declerck (1988).

There is another construction with a relative clause which is superficially similar to the transitive cleft: the extraposed relative clause modifier. Like in cleft sentences, the phrase immediately preceding the relative clause is not its antecedent (4).

- (4) De gemeente wil namelijk een breed pad aanleggen dat
 the municipality wants namely a wide path build that
 verbonden wordt met de openbare weg.
 connected becomes with the public road
Because the municipality wants to build a wide path that will be connected with the public road.

However, the transitive cleft and the regular extraposed relative clause differ on various points. First of all, while relative clause extraposition is never obligatory, the cleft clause is always extraposed: no non-extraposed variant of the *it*-cleft exists (5).

- (5) *Het die kansen voor open doel mist is immers niet de
 it that chances in front of open goal misses is after all not the
 trainer.
 coach

Furthermore, relative clause extraposition is not restricted to the pronouns *het*, *dit* and *dat*, but clefts are. Similarly, the cleft constructions is restricted to copulas, whereas relative clause extraposition is freely occurs with any verb.

We conclude that *it*-clefts with relative clauses are a construction distinct from other relative clause constructions, which calls for an analysis. In this section, we discuss the syntactic features of the different components of the construction and some previous analyses and finally present a new analysis for this construction which accounts for its characteristics and in particular the Dutch agreement facts while respecting the main principles of Dutch grammar, such as the canonical word order rules and subject-verb agreement.

2.2.2 The c-focus

The two main characteristics of the c-focus are that it is an NP and not (necessarily) predicative: proper names and pronouns can and do appear in this position. If the c-focus is an NP, it takes the complement function. We do not discuss the role of the complement of the transitive copula in this thesis. We will call it OBJ, even though we realize that it is not a regular object, e.g. it cannot passivize.

The fact that the relative clause cleft is restricted to NPs only should not be taken to mean that *all* nominals appear in clefts with relative clauses. Although the large majority of NPs combines exclusively with a relative clause, predicative nominals are not allowed in this construction. More generally, it appears that bare singular nouns combine with *dat*-clauses (6-a) instead of relative clauses (6-b). Interestingly, these bare nominals also allow for free relative extraposition (6-c), which is otherwise not possible, generally (6-f). Once combined with an article, the nominals behave like regular NPs and form relative clause clefts (6-d), but not complementizer clause clefts (6-e). Further research should be carried out to determine exactly which semantic feature is responsible for this exceptional behaviour.³

- (6) a. Het is vooral olie dat ze uitvoeren.
 it is mainly oil COMP they export
It's mainly oil that they export.
- b. *Het is vooral olie die ze uitvoeren.
 it was mainly oil REL they export
- c. Het is vooral olie wat ze uitvoeren.
 it is mainly oil FREL they export
It's mainly oil what they export.
- d. Het was vooral de olie die ervoor zorgde dat de weg
 it was mainly the oil REL for it caused that the road
 gevaarlijk werd.
 dangerous became
It was mainly the oil that caused the road to be dangerous.
- e. *Het was vooral de olie dat ervoor zorgde dat de weg
 it was mainly the oil COMP for it caused that the road
 gevaarlijk werd.
 dangerous became

³It appears that some speakers also allow non-subject forms of pronouns in complementizer clefts with non-subject “gaps” in the clause. We leave these examples aside here.

- f. ??Het was vooral de olie wat ervoor zorgde dat de weg
 it was mainly the oil FREL for it caused that the road
 gevaarlijk werd.
 dangerous became

As we saw in example (1-c), the argument structure shifts if the c-focus is a pronoun. In this case, the c-focus functions as the subject of the cleft sentence: it is in subject position and it is in the subject form. However, there are some interesting exceptions to the rule that pronouns are in subject position. Those pronouns that agree with *is* or *zijn* are occasionally found post-verbally (7-a). This is strictly ungrammatical with first and second person singular pronouns, which take the verb forms *ben* and *bent* (7-b). We do not account for this pattern here.

- (7) a. Maar het zijn wij die iets van jullie kunnen leren.
 but it are we that something from you can learn
But it's us who can learn something from you.
- b. *Maar het ben/is ik die iets van jullie kan leren.
 but it am/is I that something from you can learn

2.2.3 Agreement

In Dutch, the verb agrees with the subject in number and person. Example (1-c) shows that this is also the case in clefts: the nominative first person singular pronoun is in the sentence initial subject position and the verb shows first person singular agreement.

If we replace the pronominal c-focus in (1-c) with a full NP, the argument structure changes. The c-focus is now in object position and *het* 'it' is in subject position (1-b). Now that *het* is the subject, we expect the copula to show third person singular agreement, but surprisingly, this is not the case: if the c-focus NP is plural, the copula is plural too (1-a).

In order to account for these agreements facts, let's first look at the syntactic properties of the pronoun. The pronoun *het* in cleft constructions has often been analyzed as the expletive pronoun (Smits, 1989, for example). If this is correct, then we expect that it is impossible to replace *het* with a demonstrative, which cannot be expletive. However, we do find examples of cleft sentences with demonstratives. The examples are infrequent and mainly found in spoken Dutch, but nevertheless grammatical (8).⁴ This is similar to the German cleft construction, which also allows for a demonstrative pronoun

⁴In the following, whatever we say about the agreement features of *het* 'it' also applies to the pronouns *dat* 'that' and *dit* 'this'.

instead of the German pronoun *es* ‘it’ (Smits, 1989). Also for English, it has been claimed that so-called *th*-clefts are not impossible (Hedberg, 2000). We thus assume that *het* is not an expletive subject. The standard tests for expletiveness, such as the possibility to stress the word or to have emphatic reflexives, fail to distinguish between referential and expletive uses of *het*, as it is a weak and obligatorily stressless pronoun, but do not falsify our assumption that the pronoun is not expletive.

- (8) Goh dat is mezelf die ik hoor
 gosh that is myself that I hear
Gosh, it's me that I hear

Hedberg (2000) also advocates a non-expletive *it* in *it*-clefts. She argues that the pronoun (together with the final clause) functions as a definite description. Furthermore, Gundel (1976) argues for a non-expletive subject in cleft constructions based on data from Russian. Russian does not have expletive subjects, but it does have *eto* ‘it’ in cleft sentences.⁵ Finally, we will see below that there are remarkable similarities between the use of *het*, *dit* and *dat* in cleft sentences and their use in referential simplex copular sentences (also known as truncated clefts). Our hypothesis that the subject pronoun in NP *it*-clefts is different from the commonly assumed expletive pronoun is thus supported by previous work on cleft constructions in other languages, as well as by crosslingual and in-language data.

Secondly, we have to determine the syntactic role of the pronoun. Is *het* really the subject? Dutch has a clear canonical word order, and the pronoun is in the canonical subject position, but various arguments and adjuncts can appear in the canonical sentence initial subject position by means of topicalization. If *het* in example (1-a) is in fact the topicalized object and the plural NP is the subject, then the plural agreement on the verb would be in accordance with subject verb agreement. This analysis fails for multiple reasons. In the first place, topicalized objects must be stressed and *het* is obligatorily unstressed. Therefore, the object pronoun *het* cannot undergo topicalization. Secondly, embedded clauses do not allow topicalization. If *het* were the object, we would expect it not to show up in the subject position

⁵The relevant data are in (i). Note that the Russian construction does not contain a relativizer. Thanks to Lev Blumenfeld and Dimitry Kochenov for sharing their intuitions with me.

- (i) Eto Ivan mne pozvonil
 it Ivan me called
It was Ivan who called me

immediately following the complementizer in embedded clauses. But it turns out that the pronoun does occur in this position (9), if the c-focus is a NP. *Het* is in object position in clefts with a pronominal c-focus, just like in main clauses (10).

- (9) Ze zijn er inmiddels van doordrongen dat het de ondernemers
 they are there by now of convinced that it the producers
 zijn die de welvaart voor het volk creëren.
 that the prosperity for the people create
By now they are convinced that it is the producers that bring prosperity to the people.
- (10) Hij herkent de man en weet dat hij het is die hem
 he recognizes the man and knows that he it is that him
 binnenkort het land uit wil jagen.
 soon the country out wants chase
He recognizes the man and knows that it's him who wants to chase him out of the country shortly.

Additional evidence for the subject-hood of *het* can be found in raising constructions, where the main verb functionally controls the subject of the embedded verb. If we assume that the pronoun *het* is the subject of the cleft sentence, then it should be possible to raise it if we embed the cleft in a raising construction. And we do indeed find such raised cleft constructions (11-a). Recall that in clefts with a pronominal c-focus, *het* was the object and the c-focus was the subject. Although we did not find any examples in our corpus, it does appear possible to raise the focused pronoun when we embed (1-c) under a raising verb (11-b).

- (11) a. Het lijken vooral dit soort instellingen te zijn die in de
 it appear mostly this type organizations to be that in the
 problemen zijn geraakt.
 problems are come
It appears to be mostly this type of organization that came into trouble.
- b. Dus toen leek ik het te zijn, die stom deed
 thus then appeared I it to be who stupid did
So at that point it appeared to be me who was acting stupid.

Now that we have established that the pronoun *het* is the subject, how can we account for the agreement features of the verb? The examples (9) and (11-a) illustrate that both the embedded copula and the raising verb show plural

agreement. Following the strict subject verb agreement in Dutch, we have to conclude that *het* is plural in the examples (1-a), (9) and (11-a), similar to the analysis of *there* in phrase structure grammars (Pollard and Sag, 1994).

There is independent motivation for the existence of a plural and/or common *het/dat/dit* ‘it/that/this’. The distribution of these pronouns is not restricted to clefts and raising constructions: they also show up in other types of copular sentences, both as personal pronouns (12-a) and resumptive pronouns (12-b) (see also Rullman and Zwart (1996)). A classic discussion in Dutch linguistics deals with the question which of the constituents in sentences like (12-a) is the subject (Merckens, 1961; Bos, 1961), where the word order suggests that *dat* is the subject, but subject verb agreement suggests that *soldaten* ‘soldiers’ is the subject. It is possible to analyze the pronoun as the subject (in accordance with the Dutch word order rules) and account for the plural agreement on the verb if the pronoun has a plural value for NUM.

- (12) a. Dat zijn pas echte soldaten.
 that are now real soldiers
 Now those are real soldiers.
- b. VUT’ers, DOP’ers - dat zijn vroeg gepensioneerden en
 VUT-ers DOP-ers - that are early retired ones and
 economisch zelfstandigen.
 economical independent ones
 VUT-ers, DOP-ers, those are the early retired and the economically independent.

In these examples, the pronoun itself does not show agreement, but subject verb agreement in example (12-a) and resumptive pronoun antecedent agreement in example (12-b) indicate that the value for NUMBER on the subject is in fact plural (and GENDER is common). Based on the fact that some pronouns can have both singular and plural number and both neuter and common gender, one may think that these pronouns are simply underspecified and the finite verb has defining equations specifying its subject’s agreement features (13-a).

- (13) a. *zijn*: V (↑PRED) = ‘be-equal-to((↑SUBJ)(↑OBJ1))’
 (↑SUBJ PERS) = 3
 (↑SUBJ NUM) = pl

- b. *zijn*: V (↑PRED) = ‘be-equal-to⟨(↑SUBJ)(↑OBJ1)⟩’
 (↑SUBJ PERS) =_c 3
 (↑SUBJ NUM) =_c pl

This would explain the grammaticality of the examples (12-a) and (12-b) but not the ungrammaticality of (14-a) and (14-b), in which the use of a “non-agreeing” pronoun is not allowed.⁶ Given the entry in (13-a), nothing prevents unification of the underspecified subject pronoun and the defining subject specifications on the verb. The ungrammaticality of these examples follows naturally if we assume instead that the agreement constraints on the verb are constraining equations (13-b): the subject pronoun is not specified for number and person and therefore the constraining equations on the verb are not satisfied. But if the agreement equations are constraining, how do we explain the grammatical and apparently non-agreeing examples? To answer this question, we look at the restrictions on the use of plural/common *het*.

- (14) a. *Dat zijn pas dapper.
 that are now brave
 b. *VUT’ers, DOP’ers - dat zijn vroeg gepensioneerd en
 VUT-ers DOP-ers - that are early retired and
 economisch zelfstandig.
 economical independent

There are two ways of using the pronouns *het*, *dit* and *dat* grammatically as subjects of a simple copular clause. As we saw earlier, the pronouns may be used as singular or plural subjects of a copular clause with a nominal object (12-a), (12-b). If no nominal object is present, the pronoun *het* ‘it’, the resumptive pronoun *dat* ‘that’ and the demonstratives *dit* ‘this’ and *dat* ‘that’ are still possible, but only if they are singular and neuter (15-a). In other words: the apparent non-agreement is only possible with a nominal object.

- (15) a. Jazid laat mij het water proeven. Het is koel.
 Jazid lets me the water_{neut} taste it is cool
 Jazid lets me taste the water. It is cool.
 b. *Jazid laat mij de drankjes proeven. Het zijn koel.
 Gazid lets me the drinks taste it are cool

This pattern can be accounted for if we assume that the pronoun has optional default agreement features and agrees in number and gender with the copular

⁶The examples are fine if an explicitly plural pronoun such as *die* ‘those’ is used instead of *dat* ‘that’.

complement: if this complement is a nominal phrase, then it shares its agreement features with the pronoun, possibly causing the pronoun to be plural and/or common. If the complement is adjectival, this sharing is impossible, because the adjective is not defined for number or gender. As a result, the pronoun cannot be made plural or common and only can be used with its “default” values, neuter singular. We assume a lexical entry for *het* as in (16)⁷ and entry (13-b) for *zijn*. The specifications for number and gender are optional, so that they can be overridden by the agreement features of a nominal complement. In case the subject cannot ‘get’ agreement values from the (adjectival) predicate, it can only satisfy the constraining equations on the verb by instantiating the default value for number and person: singular neuter. This explains why the examples (14-a), (14-b) and (15-b) are out, but (12-a), (12-b) and (15-a) are ok.

(16)	<i>het</i> :	PN	(↑PRED)	‘pro’
			(↑PERS)	3
			((↑NUM)	sg)
			((↑GEN)	neut)
			(↑PRONTYPE)	cop

In addition, the pronoun has a feature PRONTYPE with value ‘cop’ (copular). This feature-value pair sets *het* ‘it’, *dat* ‘that’ and *dit* ‘this’ apart from all other pronouns. It reflects the fact that these three pronouns form a distinct class with a specific syntactic distribution and semantics (Declerck, 1988).

Subject-complement agreement in number and gender is not observed in all Dutch copular sentences. Number agreement is widespread, but there are exceptions, such as bare singular nouns which are used to predicate over plural subjects and sentences like the following example from the web (17), where the number is not shared.

(17)	dat	als	zij	mij	waren,	ze	SPF	niet	zouden	noemen.

Gender agreement across the copula exceptional in Dutch. Nouns have a fixed, lexical specification GEN, which makes it impossible to ‘adjust’ gender in order to agree with the subject. As a consequence, gender mismatches between subject and complement in copular sentences (18) are very common.

⁷The parentheses around the optional features translate to the following disjunction: (↑NUM) ∨ (↑NUM)=sg and (↑GEN) ∨ (↑GEN)=neut.

- (18) Het knelpunt is de ondoorzichtigheid.
 the bottleneck_{neuter} is the opacity_{common}
The bottleneck is the opacity.

Only some pronouns have flexible GEN specifications: demonstratives have two forms, one for neuter and one for common gender (e.g. *dat* 'that_{neuter}' vs. *die* 'that_{common}'). Here we do in fact find the expected contrast in referring expressions, as illustrated in the constructed example (19). But we saw that we get the copular pronouns *het*, *dit* and *dat* (in which the gender contrast is not observable on the surface) in sentences with nominal predicates. In addition, the common gender pronoun *die* is becoming more and more acceptable referring to neuter objects. Something similar is happening with the personal pronoun *hem* 'him', which is also used for both neuter and common objects. In short, although gender agreement within the NP (e.g. between the article and the noun, concord agreement) is very strict in Dutch, gender agreement between NPs (index agreement) is rare.

- (19) a. Ik heb mijn oude broek weggedaan. Die was
 I have my old trousers_{sg,comm} thrown-away that_{comm} was
 inmiddels te klein geworden.
 by-now too small become
I have thrown out my old trousers. By now they have become too small.
- b. *Ik heb mijn oude broek weggedaan. Dat was
 I have my old trousers_{sg,comm} thrown-away that_{neut} was
 inmiddels te klein.
 by-now too small become
- c. Ik heb mijn oude overhemd weggedaan. Dat was
 I have my old shirt_{sg,neut} thrown-away that_{neut} was
 inmiddels te klein geworden.
 by-now too small become
I have thrown out my old shirt. By now it has become too small.
- d. ?Ik heb mijn oude overhemd weggedaan. Die was
 I have my old shirt_{sg,neut} thrown-away that_{comm} was
 inmiddels te klein geworden.
 by-now too small become

One may argue that Dutch does have a general principle of subject-predicate agreement in copular sentences, but it only shows in cases where we have a pronoun. This can be modeled with a violable constraint in the Optimality Theoretic (OT) tradition. This constraint on agreement should

be outranked by a faithfulness constraint stating that lexical gender specifications should be faithfully realized: only some pronouns can satisfy both constraints and in all other cases we will get a violation of the lower ranked constraint, resulting in a gender mismatch. Alternatively, one may say that this is a peculiarity of those pronouns. This can be modeled by encoding the agreement constraint on the lexical entries of the pronouns.⁸ Our analysis is compatible with both an OT-style approach and pronoun specific functional annotations.

2.2.4 The relative clause

Clefts with a nominal *c*-focus have a final relative clause.⁹ The relativizer appears to agree in gender with the *c*-focus: *die* for common singular nouns and plurals and *dat* for neuter singular. It would nevertheless be incorrect to state that the clefted element is the antecedent, because the embedded verb does not agree in person with the *c*-focus (1-*c*), as it does in adjoined relative clauses (20).¹⁰ In section 2.2.1 it was furthermore noted that cleft clauses differ from relative clause modifiers of the predicate with respect to prosody, semantics and pragmatics.

- (20) En ik, die dit vertel ben Tina.
 And I, who this tell_{1sg} am Tina
And I, who tell this, am Tina

Alternatively, one could assume that the object is the antecedent. This gives the same results in most cases, since the object and the FOCUS usually coincide. But not if the *c*-focus is a pronoun. In that case the object (and

⁸The annotation needed to put this constraint on the copular pronoun is rather complex, as it needs to account for both argument structures: $(\uparrow\text{NUM})=((\text{SUBJ}\uparrow)\text{OBJ NUM}) \vee ((\text{OBJ}\uparrow)\text{SUBJ NUM})$ (and a similar one for gender). This constraint states that the number value of the pronoun is unified with the number value of the object if the pronoun is the subject, and NUM is unified with the number value of the subject if the pronoun is the object.

⁹We assume an analysis of relative clauses along the lines of Dalrymple (2001) and Falk (2001): the relative clause is a headless CP with the relative pronoun in SpecCP. The fronted phrase is the TOPIC of the embedded clause and the *f*-structure of the relative pronoun is the value of a feature RELPRO

¹⁰In old-Dutch and in some bible texts one can also find first person verbs in *it*-clefts. We do not account for these archaic examples here.

- (i) Ik ben het die uw overtredingen uitdelg om mijnentwil.
 I am it that your transgressions take-away_{1sg} for my-wish
It is me that takes away your transgressions because that is my wish.

thus the antecedent) is *het*. The third person agreement on the embedded verb in (1-c) would thereby be explained.

Example (21) appears to be a counterexample to this analysis: the embedded verb is plural, whereas the antecedent is *het*. Similarly, the relativizer in (22) is of common gender, while the antecedent is *het*. However, with the lexical entry proposed in (16), these examples are no longer problematic, as *het* unifies its number feature with that of the pronoun *we* ‘we’.

- (21) Wij zijn het die alle partijen bij de les moeten houden
 we are it who all parties at the lesson must hold
It's us who should make sure all parties stay focused
- (22) Hij was het ook die P.J.H. Cuypers in de arm nam [...].
 he was it too that P.J.H. Cuypers in the arm took
It was him, too, who got P.J.H. Cuypers involved.

One disadvantage of the object antecedent approach remains: its discourse function. The cleft construction is a focus construction. It focuses one element (the clefted element or c-focus, mapped to FOCUS in f-structure), while the given or backgrounded information is extraposed. Under the object antecedent analysis, the old information from the clause is analyzed as a modifier of the object, which in most cases is the c-focus. Thus, it will be part of FOCUS. This is in contradiction with it being given or background information. Furthermore, the information structure of clefts is assumed to be the same, irrespective of the syntactic category of the c-focus. But under the object antecedent analysis, the clause is part of FOCUS if the c-focus is a full NP but not if c-focus is pronominal (because the c-focus is the subject in that case). This makes the object antecedent approach an unattractive analysis.

The fact that it is difficult to find an antecedent for the relative clause, has led to the hypothesis that there is no antecedent and the relative clause is a free relative. Akmajian (1970) analyzed English clefts as pseudo-clefts that had undergone a transformation, moving the free relative to the right edge. A closely related analysis was presented for Dutch in Van der Beek (2001). There, the extraposed clause is analyzed as a free relative clause that is extraposed by means of independently motivated extraposition rules. The analysis of the final clause as a free relative accounts for the agreement facts: if the free relative is in fact the extraposed subject, then the plural free relative in (1-a) does agree with the plural verb.

An important counterargument to free relative accounts is that the form of the relative clause is not the same as a free relative: instead of the relativizers *die* and *dat* for common and neuter antecedents, free relatives use

wie and *wat* for free relatives referring to animate versus inanimate objects.¹¹ Furthermore, free relatives are always singular, with a universal or exhaustive reading, whereas the final clause of a cleft can be plural (1-a).

A second problem for the free relative analysis is that extraposition of free relatives involves expletive insertion. Both the free relative and the expletive map to the same argument function, so that the requirements of both coherence and completeness are met. As we have seen, there is reason to believe that the pronoun is in fact *not* expletive. This means that *het* has a PRED feature, which would clash with the PRED value of the extraposed subject under the free relative analysis.

The relative clause is not a modifier of the OBJ or FOCUS and it is not a free relative. That leaves two possible analyses: the antecedent of the final clause is either the SUBJ or TOPIC. The subject antecedent analysis was first suggested for English by Jespersen (1927). According to this analysis, the final clause is a relative clause that restricts the interpretation of *it*. In English, this is always both subject and topic. Jespersen developed his analysis for English and thus does not address the Dutch agreement pattern: it does not follow from this analysis that the relativizer obligatorily has the same gender as the clefted element in Dutch nor that the verb in example (1-a) should be plural. With the lexical entry for *het* presented in (16) and SUBJ-OBJ agreement in copular constructions, the agreement pattern could be accounted for. But the Jespersen analysis has the same disadvantage as the object antecedent analysis: the discourse function of the relative clause would vary.

That brings us to the analysis we propose in this chapter: the relative clause as a modifier of the topic pronoun *het*. This analysis predicts the correct NUM and GEN values if we combine it with the lexical entry for *het* discussed before. The NUM and GEN values of the pronoun unify with those of the object. The verb can now check for the appropriate values on the subject, which is either the pronominal c-focus or the topic pronoun *het* with the unified agreement features of the object. The agreement between the relativizer and the antecedent is also unproblematic under this analysis, because the antecedent *het* now has the same agreement features as the c-focus. The relative clause is always a part of TOPIC. This nicely reflects the observation that the information in the final clause of a cleft has to be given (Declerck, 1988). Our analysis resembles the analysis in Hedberg (2000), who claims that the pronoun and the relative clause function as a discontinuous definite

¹¹But note that the reference grammar of Dutch Haeseryn et al. (1997) does allow *die* and *dat* as the heads of free relatives, although the non-cleft examples are marginal. In addition, the dictionary of Dutch from 1500-1976 does list them as possible heads of free relatives (de Vries et al. (1882-1998), column 2517-2518)

description. But in her syntactic analysis of the construction, she analyzes the clause as adjoined to the focused phrase. The LFG framework facilitates an account of *it*-clefts in Dutch in which the pronoun and the clause form a unit both semantically and syntactically: the two components map to same f-structure even if they are discontinuous on the level of c-structure. This analysis is formalized in the next section. The construction is treated as a specific focus construction, distinct from regular relative clause extraposition. The idiosyncratic properties of the construction (see also section 2.2.3) can thus be dealt with.

Note, finally, that Dutch has another construction in which an obligatorily clause final relative clause modifies a pronoun. These are instances of the quantitative use of the R-pronoun *er*, where the quantitative element is left out (23). Like in clefts, the pronoun and the extraposed relative clause do not form a constituent on c-structure.

- (23) Maar er zijn er ook die het met achthonderd dollar in
 but there are R-pron also that it with eight hundred dollar in
 de maand moeten doen.
 the month must do
But there are also people who must do with 800 dollars a month.

2.2.5 Formalization

We have argued that the pronoun *het* ‘it’ has a lexical entry with default agreement values. The transitive Dutch *it*-cleft consists of this pronoun, a second nominal argument (the c-focus) and a relative clause. The antecedent of the relative clause is the topic pronoun *het*.

The different parts of the analysis are combined in the c-structure rules in figure 2.3 on page 49. The rules are for main clause clefts. Although the c-structure rules for subordinate clauses are different, the idea is the same: two nominal arguments and a relative clause on the right edge. It is this relative clause that carries the construction specific f-structure specifications for focus on the clefted element and the pronoun *het* with discourse function TOPIC in either subject or object position, bearing a feature ADJ that is filled by the final clause as a whole.¹² An example c-structure is given in fig. 2.1, the corresponding f-structure in fig. 2.2.

Like in regular relative clauses, the relative pronoun in the final clause can be embedded in a PP (24). These examples are automatically accounted

¹²The concept of a sentence final CP that maps to the ADJ of the non-expletive pronoun *it* is also found in Berman’s analysis of extraposed argument clauses in German (Berman, 2001).

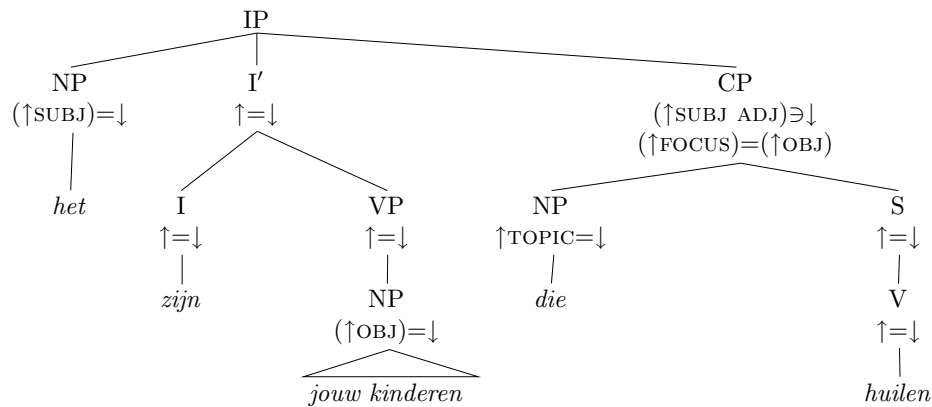


Figure 2.1: C-structure for *Het zijn jouw kinderen die huilen*

for by the regular relative clause rules.

- (24) Het zijn dat soort reacties waardoor de military-ruiters
 it are this kind reactions through-which the military riders
 zich onbegrepen voelen.
 themselves misunderstood feel
It is this type of responses that make the military riders feel misunder-
derstood.

The c-structure rules in figure 2.3 show that the transitive *it*-cleft has various construction specific features that have to be stipulated in the c-structure rule: the relative clause does not form a \bar{N} with its antecedent, the relative clause is obligatorily at the right edge and the TOPIC has to be of a particular pronoun type. On the other hand, we used the independently motivated c-structure rules for transitive sentences and specialized them in order also to cover cleft sentences. The only component that was added is the optional relative clause with all the construction specific information.¹³

This analysis leaves the canonical Dutch word order intact: the canonical subject position, filled by *het*, is associated with the grammatical subject function. At the same time, it meets the requirement of subject verb agreement: the pronoun *het* is in fact plural, since it unifies its AGR values with those of the object. This unification also predicts the observed pattern of agreement between the relativizer and the pronoun.

We do not account for the distribution of the two argument structures

¹³The rules for the transitive cleft can be merged with the general rules for transitive clauses by adding the CP optionally to the general IP rule.

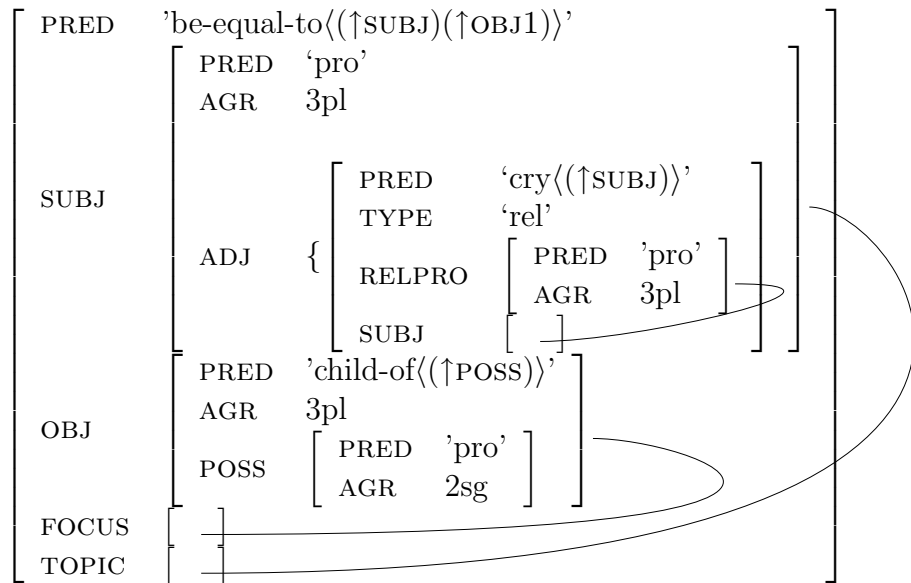


Figure 2.2: F-structure for *Het zijn jouw kinderen die huilen*

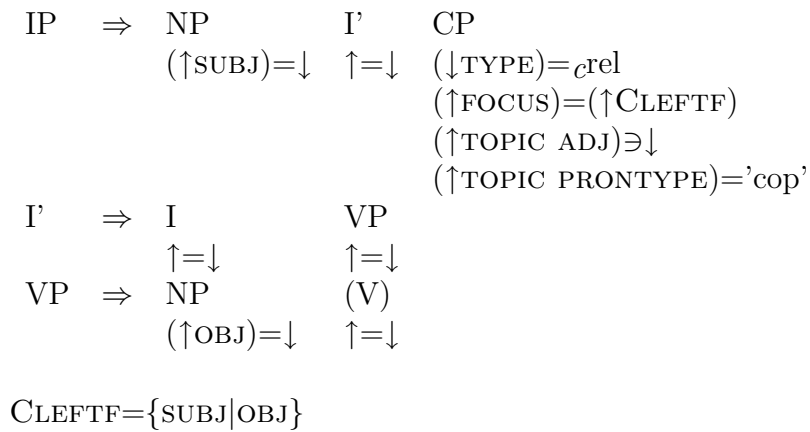


Figure 2.3: C-structure rules for nominal clefts

of the transitive cleft in the c-structure rules. The rules in figure 2.3 generate both argument structures for both pronouns and full NPs, even though focused pronouns are virtually always realized as subjects in a cleft construction, and focused NPs are realized as objects in a cleft construction. It is assumed that general constraints penalize copular object pronouns, and focused pronouns in particular. Thus candidates like (25) are excluded.¹⁴

- (25) a. *omdat het HEM is die huilt
 because it him is who cries
 b. *omdat jouw ZOON het is die huilt
 because your son it is who cries

This assumption that the argument structure in clefts is an effect of a more general mechanism is supported by the fact that the same effects can be observed in other copular sentences. Haeseryn et al. (1997) list 3 copular constructions in which pronouns can function as a complement (26). In all cases, the subject is a pronoun, too. Sentences with a full NP subject and a pronominal complement are ungrammatical (27). Apparently, pronouns are realized as subjects whenever possible.¹⁵

Only if both functions are realized by pronouns is a pronominal complement acceptable. The only surprising example is (26-c), where the subject is *het* and the complement a personal pronoun. After all, this argument structure is out in cleft sentences. Coppens (1996) noted that the same string with a neutral stress pattern is ungrammatical (28-a). He accounts for the ungrammaticality of the example based on the assumption that the copular complement is thematically associated with the subject, but receives the non-subject case because of its position. This is no problem for nouns or the pronouns *het*, *dat* and *dit*, which do not carry casemarking, but it is a problem for personal pronouns, the only category in Dutch which does show case marking. He does not account for the contrast between (28-a) and (26-c). For cleft sentences, it is important to note that the stress pattern in (26-c) is not available: stress is on the c-focus.

¹⁴In fact a number of examples like (i) can be found on internet. This shows that the constraint is not categorical. In the next chapter we investigate how one may account for such non-categorical distinctions.

- (i) Ik denk dat het hem is die ik bedoelde
 I think that it him is that I meant
 I think it was him that I meant.

¹⁵This assumes that the relation between the two copular arguments in these constructions is symmetrical.

- (26) a. Als ik jou was ...
 if I you were ...
 If I were you ...
- b. Hij is 'm.
 he is him_{reduced}
 He is it.
- c. Het IS 'm
 it IS him_{reduced}
 It IS him.
- (27) *dat kandidaat A jou is.
 that candidate A you is
- (28) a. *Het is hij/hem.
 it is he/him
- b. Hij is het.
 he is it
 It's him.

It appears appropriate to treat the argument structure differences in clefts and other copular constructions as part of a yet more general distributional phenomenon. If we compare the argument functions of *het*, personal pronouns and full NPs, we find striking differences. For all categories, the subject function is most frequent, but this preference is much stronger for pronouns than for NPs. Looking at subjects, direct objects and predicative complements only, we find for NPs the following distribution in the Alpino Treebank: 65% subject, 29% direct object and 6% predicative complement. For *het*, the distribution is 77% subjects vs. 22% direct objects and 1% predicative complements.¹⁶ Finally, of the 2075 relevant personal pronouns, 94% had the subject function and 6% had the direct object function; only one personal pronoun functioned (grammatically) as a predicative complement (29). Although it is technically possible to constrain subject foci to personal pronouns with functional annotations, we assume that our argument order variation is a direct consequence of these more general distributional phenomena.

- (29) Dat was 'm dan, de Puskas van het Poolse voetbal.
 that was him then the Puskas of the Polish soccer
 So that is him, the Puskas of Polish soccer.

¹⁶Discarding sentences with expletive, preliminary subjects and extraposed sentential subjects.

2.3 Intransitive Clefts

So far, we only looked at clefts with a final relative clause. But the clefts in (1-d) and (1-e) do not contain a relative clause. In the next section we discuss this and many other differences between the clefts in (1-a)-(1-c) (transitive clefts) on the one hand, and (1-d)-(1-e) (intransitive clefts) on the other hand.

2.3.1 Differences between transitive and intransitive clefts

The clause The first difference between both types of clefts is the final clause. While transitive clefts have final relative clauses, this second type of cleft has a subordinate final clause headed by the complementizer *dat*. Although this complementizer is homonymous to the neuter singular relativizer, we know that it is in fact a complementizer because there are no neuter singular “traces” in the clause.

The claim that intransitive clefts have a final complementizer clause is not universally agreed upon. Smits (1989) argues that the word *dat* introducing the clause is of a special syntactic category called *relative particle*, which introduces a relative clause. This particle is not only used in intransitive clefts, Smits claims, but also in a specific type of relative clause, namely one that modifies a temporal expression (30).

- (30) Hamills rol [...] in Jay and Silent Bob Strike Back is de eerste
 Hamil’s role [...] in Jay and Silent Bob Strike Back is the first
 keer dat hij de draak steekt met Star Wars.
 time that he makes fun of Star Wars
*Hamil’s role [...] in Jay and Silent Bob Strike Back is the first time
 that he makes fun of Star Wars.*

This is in line with Smits’ (informal) definition of relative clauses as “any construction part of which is a subclause that modifies an expression external to that subclause [...]”. It furthermore allows for the generalization that any cleft has a relative clause. But there are a number of problems with this analysis. First of all, the assumption of a relative particle raises the question why the usage of this particle in clefts is so much broader than in regular relative clauses. Secondly, it is unclear what the antecedent is of the particle in cleft sentences, especially since the clause does not seem to modify any element in the c-focus. Furthermore, one wonders why this relative particle cannot function as an argument in the embedded clause (31), just like relative pronouns. And finally, we will see later on that we do find intransitive cleft

sentences (but not modifiers of temporal expressions) with argument ‘traces’ in the embedded clause. It is unclear how Smits (1989) would account for this contrast.

- (31) *De eerste keer dat ik mij goed herinner was op 5 februari.
 the first time that I REFL well remember was on 5 February

We assume that a word which looks and behaves like a complementizer is in fact a complementizer and introduces a complementizer clause. We thus have to distinguish between clefts with relative clauses and clefts with complementizer clauses.

C-focus Secondly, there are differences with respect to the categories that may be focused. Transitive *it*-clefts only focus nominals, but complementizer clefts may focus a wide range of categories, including PPs (1-d), CPs (1-e) and AdvPs (32-a). *it*-clefts with APs (32-b) have been reported grammatical in the literature Smits (1989), but intuitions differ from one speaker to another and no corpus examples were found. In section 2.2.2, we furthermore saw that a restricted set of nominals occurs not in the relative clause construction, but in the complementizer construction.

- (32) a. Het is daar dat de verveling toeslaat.
 it is there that the boredom attacks
 It's there that boredom attacks.
 b. ?Het is rood, dat hij zijn kamer verft
 it is red that he his room paints
 It's red that he paints his room

The pronoun Another difference between the two constructions is the fact that the pronoun *het* cannot be replaced by a demonstrative (33-a), as it could in transitive clefts. This suggests it is an expletive, as members of like categories are otherwise generally interchangeable.

- (33) a. *Dat is daar dat de verveling toeslaat.
 that is there that the boredom attacks
 b. *HET is daar dat de verveling toeslaat.
 IT is there that boredom attacks
 c. *Hetzelf is daar dat de verveling toeslaat.
 itself is there that the boredom attacks

In section 2.2 it was already shown that other tests for expletiveness fail to make a clear distinction between the expletive *het* and the non-expletive but

obligatorily unstressed pronoun *het*, but there is some further evidence that the pronouns in the two constructions are dissimilar. Recall that Gundel (1976) argued for a non-expletive cleft subject based on data from Russian, which does not have expletive subjects, but does have *eto* ‘it’ in cleft sentences. This appears to contradict our assumption that *het* is an expletive in Dutch complementizer clefts, but the Russian examples all have a nominal focus. It turns out that the *eto*-cleft is in fact only possible with a nominal phrase in focus: different constructions are used to focus PPs or complementizer phrases. Hedberg (2000), which also argued against expletive cleft subjects, explicitly restricts the argument to NP *it*-clefts only. As the large majority of Dutch NP clefts are of the transitive type, this supports our hypothesis of a non-expletive subjects in transitive clefts, without contradicting the expletive status of the subjects in intransitive clefts.

The copula The two types of clefts also differ with respect to the verb. Relative clause clefts may use another copula instead of *zijn* (34-a), while this appears not to be possible for clefts with complementizer clauses (34-b).¹⁷

- (34) a. Toch bleken het uitgerekend de Democraten die [...] het
 Yet appear it calculated the Democrats that [...] the
 meest op hadden met de watersnood.
 most on had with the flooding
*And yet, of all parties, it turned out to be the Democrats who
 cared most about the flooding.*
- b. *Toch bleek het op Democratisch initiatief, dat er
 Yet appeared it on Democratic initiative that there
 steun voor de watersnood kwam.
 support for the flooding came

Argument structure The transitive cleft has two arguments, a subject and an object. This object is usually the c-focus. It is difficult to analyze the complementizer clefts in the same way. The c-focus of an intransitive cleft would make a very unusual object, since it is almost never nominal. And semantically, the construction does not resemble a transitive sentence either: in contrast to the relative clause clefts, the complementizer clefts (35-a) cannot be paraphrased as canonical specificational sentences (35-b), even if we transformed the *that*-clause into a locational free relative (35-c). The best paraphrase would be the simplex sentence in (35-d).

¹⁷Thanks to Frank Van Eynde for pointing this out to me.

- (35) a. Het was in Polen dat het eerste vrije vakverbond onder het
it was in Poland that the first free union under the
communisme werd opgericht.
communism was founded
*It was in Poland that the first free union under communism was
founded.*
- b. *Dat het eerste vrije vakverbond onder het communisme werd
that the first free union under the communism was
opgericht was in Polen.
founded was in Poland
- c. *Waar het eerste vrije vakverbond onder het communisme
Where the first free union under the communism
werd opgericht was in Polen.
was founded was in Poland
- d. In Polen werd het eerste vrije vakverbond onder het
in Poland was the first free union under the
communisme opgericht.
communism founded
The first free union under communism was founded in Poland.

Finally, the transitive cleft has a variable argument structure and both *is* ‘is’ and *zijn* ‘are’ occur in the matrix clause. In the intransitive cleft, the pronoun is always in subject position and the verb is always singular.

We conclude that the construction that is generally called the *it*-cleft construction consists of two distinct constructions, at least in Dutch: one with a relative clause, and one with a complementizer clause. A similar claim has been made for English clefts by Pinkham and Hankamer (1975). They argue that in English, non-nominal clefts must be derived from simplex sentences such as (35-d) by means of extraction of the focus, creation of the matrix copular construction, and extraposition of the original sentence. Nominal clefts on the other hand are derivationally ambiguous: they may be derived from a simplex sentence in a similar way as non-nominal clefts, or alternatively the copular construction may be base generated with a headless clausal subject, which is then relativized and extraposed, leaving behind an (expletive) pronoun *it*. The two proposals share certain features, e.g. only nominals combine with relative clauses in clefts. But the data, arguments and analyses differ greatly. Pinkham and Hankamer (1975) do not say much about agreement, as agreement in English clefts is always third person singular. The analysis focuses primarily on the status of the verb *to be*, which

is base generated in one derivation but not in the other. From this, they predict facts about connectivity, reflexivization and negation, which do not translate to the Dutch construction. The dual derivation analysis does not look at the differences with respect to the pronoun, the verbs that can be used, or the semantics. Furthermore, the analysis presented in this thesis does not assume the nominal cleft to be derivationally ambiguous, while this is an important feature of the analysis in Pinkham and Hankamer (1975).

Before turning to the analysis of this second type of cleft, we first discuss the differences between complementizer clefts and other complementizer clause constructions.

2.3.2 Differences between intransitive clefts and other complementizer constructions

We have argued that the sentences which are normally considered to be of the same kind (i.e. *it*-clefts) are in fact two distinct constructions and we have proposed an analysis for the first: the typical cleft construction with a nominal c-focus was analyzed as an instance of the transitive (specificational) copula. The question arises what sort of construction the other sentences are. They consist of an expletive subject, the verb *to be* and an extraposed complementizer clause. Two types of sentences contain these same elements: sentences with *that*-clause modifiers and sentences with extraposed clausal subjects, and one may ask if our ‘clefts’ may be grouped under one of them.

That-clause modifiers are clauses headed by the complementizer *dat* that modify a preceding constituent, usually a nominal temporal expression (36-a). In section 2.3.1 we already saw an example of this construction (30). Often, the NP and *that*-clause co-occur with a subject *het* and a copula, as in example (36) below.

- (36) a. Het is de eerste keer dat een minister op deze manier
 it is the first time that a minister on this way
 ingrijpt.
 intervenes
It is the first time a minister intervenes in this way.
- b. Het is de eerste keer dat een minister op DEZE manier
 it is the first time that a minister on THIS way
 ingrijpt, maar hij heeft vaak genoeg op andere wijze
 intervenes but he has often enough on different manner
 zijn macht doen gelden.
 his power let count

It is the first time the minister intervenes in THIS way, but he has exercised his power in other ways often enough.

Although the resulting structure is similar to our clefts, there are important differences. The pronominal subject in sentences with *that*-clause modifiers is referential: it refers to a situation or an event. As a consequence, it can be replaced by a full NP, as we already saw in section 2.3.1 (30). This is impossible in intransitive clefts.

Secondly, the *that*-clause modifies the preceding phrase and forms a constituent with it. Thus, example (37-a) is a grammatical NP, but (37-b) is not a constituent nor is the nominal example (37-c).

- (37) a. de eerste keer dat een minister op deze manier ingrijpt
 the first time that a minister on this way intervenes
the first time a minister intervenes in this way
- b. *in Polen dat het eerste vrije vakverbond werd opgericht
 in Poland that the first free union was founded
- c. *olie dat ze uitvoeren
 oil that they export

Furthermore, the focus of the sentence may be inside the modifying *that*-clause (36-b). In relativizer clefts, the subordinate clause is always given and never focused: the focus is on the phrase in the object position.

Another construction that is similar to our cleft examples is clausal subject extraposition. Certain predicative adjectives and nouns allow extraposition of the subject *that*-clause (38).

- (38) a. Het is duidelijk dat het met Abu Ammar is gedaan.
 it is clear that it with Abu Ammar is over
It is clear that Abu Ammar is over.
- b. Het is van wezenlijk belang dat we nu ingrijpen.
 it is of real importance that we now intervene
It is of great importance that we intervene now.

These extrapositions are analyzed as transitive copular constructions. The subject is a preliminary, expletive subject and the clause is the ‘real’ subject. Both map to the SUBJ function. Like complementizer clefts, the construction consist of an expletive, a non-verbal complement (usually an adjective or a PP) and a *that*-clause. We nevertheless argue that the two structures are very different from one another and should not be analyzed in the same way.

Our first argument is that the predicate (the adjective or PP) in these extraposition constructions predicates a property of the whole proposition,

whereas the focused phrase in an intransitive cleft specifies a focus within the extraposed clause. Compare on the one hand the intransitive cleft (35-a) and its paraphrase with the focus *in situ* as a modifier of the event (35-d), and on the other hand the propositional predication in (39-a) and the constructed *in situ* variant (39-b), which means something completely different. This difference is reflected by the fact that we do find adverbs in focus in intransitive clefts (32-a), but not in *that*-clause extraction. Similarly, propositional predications may be non-extraposed and without *het* (39-c), but this is not possible for clefts (35-a)-(35-b).

- (39) a. Het is goed dat we de grenzen vaststellen.
 it is good that we the limits determine
It is good that we determine the limits
- b. We stellen de grenzen goed vast
 we determine the limits good
We determine the limits well.
- c. Dat we de grenzen vaststellen is goed
 that we the limits determine is good
It is good that we determine the limits

Secondly, *it*-clefts only occur with *zijn*, but we find occurrences of *that*-clause extrapositions with verbs other than *zijn* ‘to be’ (40)-(41).

- (40) Maar het lijkt evident, dat Endel hem ook hard nodig heeft.
 but it seems evident that Endel him also hard need have
But it seems clear that Endel seriously needs him as well.
- (41) Ik vind het verstandig dat Inge dit toernooi laat schieten.
 I find it wise that Inge this tournament let shoot
I think it's wise that Inge skips this tournament.

Thirdly, the final clause in a complementizer cleft is always headed by *dat* ‘that’, but this is not the case for extraposed clausal subjects. For some of the adjectives that allow for *that*-clause extraposition we also find examples with *of*-clause extraposition (42) or VP extraposition (43). We conclude that cleft sentences with extraposed complementizer clauses are different from adjectives with extraposed clausal subjects.

- (42) Het is onduidelijk of de stijging iets te maken heeft
 it is unclear whether de increase something to make has
 met het nieuwe honderdje.
 with the new hundred-diminutive

It is unclear whether the increase has anything to do with the new one hundred guilder bill.

- (43) Het is goed om kritiek te krijgen van De Nederlandsche Bank
 it is good to critique to receive from The Dutch Bank
It is good to receive comments from The Dutch Bank

2.3.3 The intransitive analysis

Complementizer clefts have to be distinguished from relativizer clefts and from other constructions with extraposed complementizer clauses. In this section we analyze them as a separate construction, based on the intransitive (existential) copula.

The complementizer cleft consists of the copula and three components: the pronoun *het*, the c-focus and the final clause. We argued in section 2.3.1 that *het* is the expletive subject in this construction, among other things because it cannot be replaced by any other nominal.

The second constituent is the c-focus. It is unlikely to be the object of the copula, because it cannot be an NP. But if it is not OBJ, then what is it? Transformationalists analyzed it as a phrase that is moved out of the final clause (Pinkham and Hankamer, 1975; Emonds, 1976). They thus derived complementizer clefts (35-a) from canonical sentences (35-d). Pollard and Sag (1994) implemented the same idea in a non-transformational way. They assume a special lexical entry for *be* for clefts. This lexical entry has three elements on the subcategorization list: *het*, an XP and a complementizer clause with that XP on slash.

If the c-focus originates in the final clause, we expect to find “traces” of it in the clause. If the extracted material is an adjunct, it is impossible to tell whether something is missing in the clause. But if it is an argument, than we should find an embedded verb which lacks one of its argument. We do indeed find examples like (44), where the c-focus realizes an argument function of the embedded verb.

- (44) a. Het is van de onwetendheid van de mensen dat ik het moet
 it is of the ignorance of the people that I it must
 hebben.
 have
It is on the ignorance of the people that I depend.
 b. Het is in deze jeugdlektuur dat het duivelse en helse
 it is in this child-literature that the devilish and hellish

zich voor het laatst manifesteren.
 self for the last manifest
*It is in this children's literature that devilish and hellish things
 last manifest themselves.*

How does this clause fit in the argument structure? The complementizer clause is not in a canonical argument position, but in a sentence final position for extraposed constituents. It is not the object of the transitive copula, because the transitive copula needs two referential, non-expletive arguments (and we already showed that the *c*-focus is an extracted part of the clause and the subject is expletive). This is illustrated in (45). Here we repeated (35-a), but we undid the extraction of the *c*-focus out of the clause, so that it is a regular *that*-clause object. The sentence is syntactically marked, but even if one judges the sentence well formed (e.g. as an answer to the question what the main point of his lecture was), its meaning is different from the meaning of the cleft sentence, because the pronoun is interpreted referentially instead of as an expletive.

- (45) ?Het was dat het eerste vrije vakverbond onder het communisme
 it was that the first free union under the communism
 in Polen werd opgericht.
 in Poland was founded

Alternatively, we analyze the copula in the complementizer clause cleft as the intransitive copula. *Het* is in subject position and maps onto the SUBJ *f*-structure as dictated by the word order rules for Dutch. It does not contribute anything to the *f*-structure besides third person singular agreement values, because it is an expletive subject. The complementizer clause (with extraposed *c*-focus) is then mapped to the same SUBJ slot. This does not lead to a clash with *het*, because it unifies with the only features of the pronoun, the AGR features. Like the expletive pronoun, the complementizer clause is always third person singular, as illustrated in sentences with CPs in canonical subject position (46).

- (46) Dat we gewonnen hebben is nog niet zeker
 that we won have is still not certain
It is not certain yet that we have won

This gives us a total of three *c*-structure nodes associated with the SUBJ *f*-structure slot: *het*, the complementizer clause and the clefted element, which is extracted from the CP. An example *c*-structure is shown in fig. 2.4. The corresponding *f*-structure is in fig. 2.5.

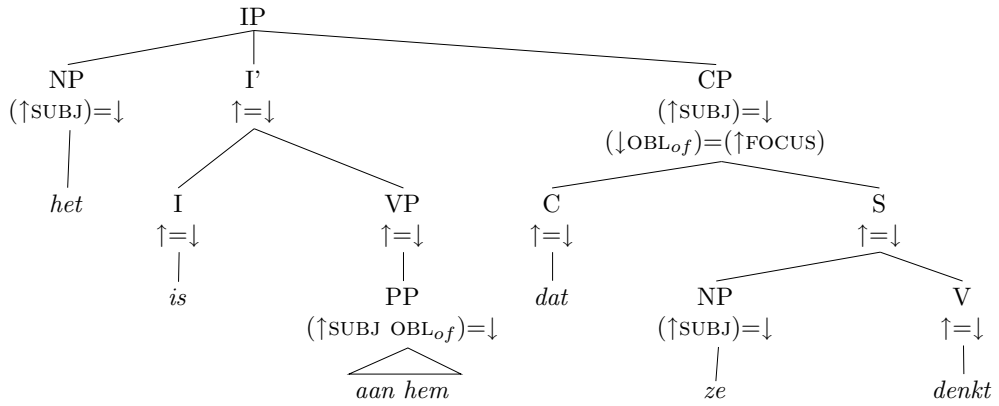


Figure 2.4: c-structure for *Het is aan hem dat ze denkt*

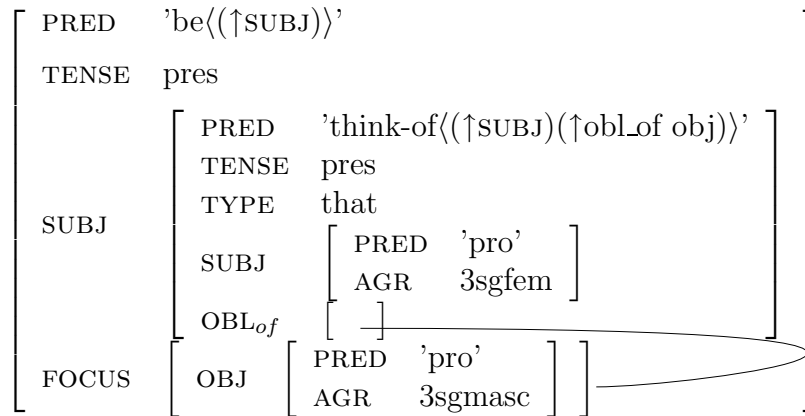


Figure 2.5: f-structure for *Het is aan hem dat ze denkt*

2.3.4 Formalization

The c-structure rules for intransitive clefts are given in figure 2.6.¹⁸ The rules also account for sentences with embedded “gaps” in the complementizer phrase, as in the constructed example (47). The clefted element is situated in the canonical object position inside the VP. This is in contrast with analyses that assume the clefted element and the final clause to be one constituent Merchant (1998); Rizzi (1997). However, this assumption does not hold for Dutch clefts, since the verb cluster obligatorily follows the clefted element and thus separates the two phrases, as in the constructed example below (48). Furthermore, the NP plus *that*-clause cannot be topicalized, as one would expect if they formed a constituent.

- (47) Het was in Polen dat ze dacht dat het eerste vrije
 it was in Poland that she thought that the first free
 vakverbond onder het communisme was opgericht.
 union under the communism was founded
It was in Poland that she thought the first free union under communism was founded
- (48) Het moet in Polen geweest zijn dat het eerste vrije vakverbond
 it must in Poland been be that the first free union
 onder het communisme was opgericht.
 under the communism was founded
It must have been in Poland that the first free union under communism was founded

The rules do not specify the NP in the canonical subject position. This is not necessary, because the expletive *het* is the only NP that would not lead to a clash in that position: every other NP has a PRED, which cannot possibly unify with the PRED of the complementizer clause because of functional uniqueness. We did not specify the argument function of the c-focus either, which means that the complementizer clause has to be instantiated to determine the syntactic function of the c-focus (or the coherence principle is violated).¹⁹

¹⁸The constraint $(\downarrow\text{CPATH})=(\uparrow\text{FOCUS})$ expands to $(\downarrow\text{XCOMP*OBL}_\theta)=(\uparrow\text{FOCUS}) \vee (\downarrow\text{XCOMP*OBL}_\theta)\ni(\uparrow\text{FOCUS})$

¹⁹The definition of CleftP should be expanded to allow for the restricted set of nominals that appears in intransitive clefts. As we have not identified the constraints on the occurrence of nominals in this construction yet, and generally allowing NPs in this construction would lead to massive overgeneration, we decided to leave out NPs for now. Additionally, for speakers who judge examples with an adjectival c-focus (32-b) grammatical, CleftP and CPath should be expanded accordingly.

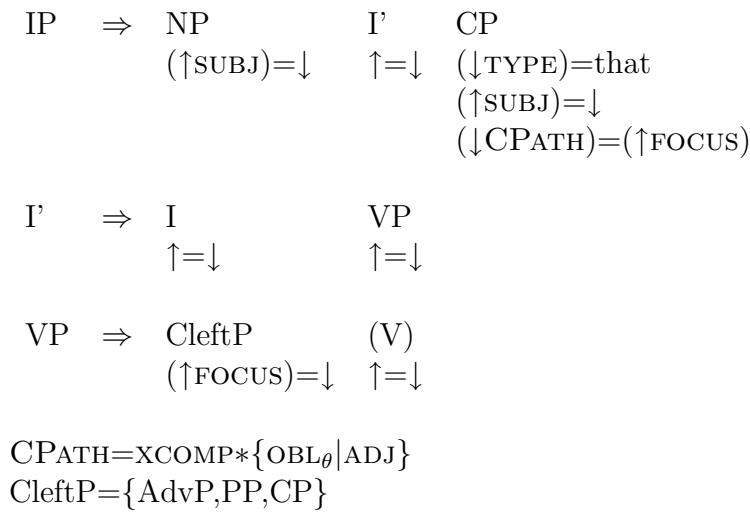


Figure 2.6: C-structure rules for non-nominal clefts

In section 2.3.1 we stated that the closest paraphrase of the intransitive cleft (35-a) is the simplex sentence (35-d). This is in line with the analysis presented in this section, which specifies the meaning of the cleft sentence to be the existential assertion of the simplex sentence (with focus on one particular constituent). Furthermore, the observation that transitive, but not intransitive clefts exist with other copulas than *zijn* follows automatically from our analysis: while the transitive copula *zijn* forms a natural class with other transitive copulas such as *lijken* and *blijken*, the latter do not have the intransitive, existential use that *zijn* has.

Note that the intransitive analysis would be inappropriate for the *it*-clefts with a relative clause, which we analyzed as transitive copular sentences. The intransitive analysis depends on the subject being expletive and we showed that this is not the case in relative clause clefts. Furthermore, the relative clause cannot independently function as an argument; it always needs an antecedent (unless it is a free relative). The two distinct analyses are furthermore motivated by the different semantics, informally illustrated by the different paraphrases. The two have in common that given information is extraposed to focus new information.

2.4 Conclusion

We have accounted for the syntactic differences between various realizations of the cleft construction by splitting up the data into two types of *it*-clefts: those with a final relative clause, and those with a final complementizer clause. We argued that the latter is a distinct construction, different both from the NP cleft and other constructions with superficially similar structures.

We analyzed the *it*-cleft with a nominal c-focus and a relative clause as an instance of the transitive copula and those with a complementizer clause as an instance of the intransitive verb *zijn* ‘to be’. We argued for a lexical entry for *het* ‘it’, *dat* ‘that’ and *dit* ‘this’ with optional agreement features. These account for the apparent lack of agreement in copular sentences with two NP arguments. Since the transitive cleft is an instance of such a sentence, we also accounted for the subject verb agreement pattern in this type of cleft. In addition, we accounted for the agreement on the embedded verb and agreement between the relative pronoun and the antecedent by analyzing it as a modifier of the TOPIC. This also explains the often observed givenness constraint in clefts: all the information in the clause has to be given or background information.

In the second type of cleft, the intransitive cleft, all phrases are associated with the subject function of the verb *zijn*: the c-focus is analyzed as an extracted constituent of the complementizer clause and both the CP and the expletive pronoun in subject position are unified with the subject function.

We did not have to stipulate construction-specific lexical entries for the copula, as clefts were analyzed as instances of the regular transitive and intransitive uses of *zijn* ‘to be’. All cleft-specific information was specified on the added, cleft-specific components in the c-structure rules. Both types of cleft involve extraposition of given information in order to focus new information.

We did not discuss the properties of the NP complement of the copula in the transitive cleft. It is clearly different from regular objects, for example in that it cannot passivize. But is different from predicative complements too, first of all in that it doesn’t have to be predicative: proper names are allowed too. In fact, we saw that purely predicative (bare) nominals are even excluded. A more precise definition of the constraints on the nominal complement and their characteristics was left for future research.

The fact that only pronominal c-foci appear in subject position was attributed to a more general phenomenon that pronouns have a strong preference for the subject function, much stronger than other syntactic categories. We expect that with further investigation along the lines of the research described

in the next chapter, a more detailed account for this particular distribution could be given.

