The Dutch it-cleft constructions

Leonoor van der Beek

Rijksuniversiteit Groningen

vdbeek@let.rug.nl

Proceedings of the LFG03 Conference

University at Albany, State University of New York

Miriam Butt and Tracy Holloway King (Editors)

2003

CSLI Publications

http://csli-publications.stanford.edu/

Abstract

In this paper I argue that the Dutch it-cleft is in fact two constructions: one with the transitive copula and a nominal focus and the other with the intransitive copula and a non-nominal focus. The analysis of transitive clefts as copular sentences with a discontinuous topic accounts for the puzzling agreement without violating the generally assumed word order rules nor subject verb agreement.

1 Introduction [1]

Dutch it-clefts are a puzzling construction. They consist of the same basic elements as English clefts—the pronoun *het* (it), a copula, the focused phrase and a final clause—but agreement is different: if the focus is plural, then the copula is plural too, even though the subject is *het* (it) (1a). 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 focus is a pronoun: if the focus is a full noun phrase, *het* is in the preverbal subject position and the focus in the post-verbal object position (1a-1b). But if the c-focus is pronominal, then it is in subject position and *het* is in object position (1c).

A second challenge for the analysis of Dutch clefts is the difference between the sentences with a nominal focus (1a-1c) and those with a non-nominal focus (1d-1e). In the first three examples, the final clause is a relative clause, headed by either the plural/common relativizer *die* (1a-1c) or the singular neuter relativizer *dat* (1b). In the examples (1d-1e) the final clause is a complementizer clause headed by the complementizer (*dat*).[2]

(1) a. Het zijn jouw kinderen, die(pl) huilen
it are your children who(pl) cry
It is your children who cry

b. Het is jouw kind(sg_neut), dat(sg_neut) huilt
it is your(sg) child(sg_neut) who(sg_neut) cries
It is your child who cries

c. Ik ben het, die(comm) dit doet
I am it who(comm) this does
It is me who does this

d. Het was in Amsterdam, dat ik hem voor het eerst ontmoette
it was in Amsterdam that I him for the first time met
It was in Amsterdam that I first met him

[1]This research is part of the Pionier Project *Algorithms for Linguistic Processing*, which is funded by the Dutch Organization for Scientific Research (NWO). I received additional funding for this research from NWO, the School of Behavioral and Cognitive Neurosciences (BCN) and the Center for Language and Cognition Groningen (CLCG). I would like to thank Joan Bresnan for comments, discussion and support. Thanks also to Gertjan van Noord, Gosse Bouma, Ron Kaplan and Gerlof Bouma for comments. Any errors are my own.

[2]The 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.

e. Het is omdat ik moe ben, dat ik thuisblijf
it is because I tired am that I home-stay
It is because I am tired that I stay at home

I will argue 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) copula for clefting other syntactic categories. I will account for the agreement patterns in (1) without violating the generally assumed fixed word order rules for Dutch nor subject verb agreement. In addition, I will show how both the argument structures in (1b) and in (1c) can be generated by one set of rules.

I do not discuss the status or characteristics of the complement of the transitive copula in this paper, nor do I say anything about the different types of uses of the copula (e.g. predicative, specificational, existential). I will call the complement OBJ, even though I realize that it is not a regular object, e.g. it cannot passivize. The two main characteristics of the non-subject cleft argument are that it is an NP and not (necessarily) predicative: proper names and pronouns can appear in this position. The predicative use of bare nouns (2) is ungrammatical in clefts. For a discussion on copular complements and different usages of the copula, see (Declerck, 1988; Higgins, 1976).

(2) * Het is dokter die hij worden wil
it is doctor that he become wants

Section 2 presents an analysis of transitive clefts and section 3 proceeds with an analysis of intransitive clefts. In section (4) I discuss some open ends and possible applications of the mechanism presented below to other syntactic constructions.

2 Transitive clefts

The first type of cleft has a final relative clause and a nominal focus, which is either a pronoun or a semantic NP (1a-1c). 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 focus.

2.1 Agreement

In Dutch, the verb agrees with the subject in number and person. Example (1c) 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 the pronominal focus in (1c) is replaced with a semantic NP, the argument structure changes. The focused constituent is now in object position and *het* (it) is in subject position (1b). 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 focus NP is plural, the copula is plural too (1a).

Note that the pronoun *het* in the cleft construction has been analyzed as the expletive pronoun (Smits, 1989). However, in clefts with a nominal focus the pronoun *het* can be replaced by the demonstrative pronoun *dat* (that) or *dit* (this), which are never expletive (3).[3] This is similar to the German cleft construction, which also allows for a demonstrative pronoun instead of the German pronoun *es* (it) (Smits, 1989).

(3) Zijn dat konijnen die daar lopen?
are that rabbits that(pl) there walk?
Are those rabbits walking over there?

Is *het* really the subject? The word order in Dutch is relatively fixed, but main clauses do allow for subject object inversion. If the plural NP in example (1a) is in fact the inverted subject and *het* is the inverted object, than the plural agreement on the verb would be in accordance with subject verb agreement. This analysis fails for multiple reasons. In the first place, inverted objects must be stressed and *het* is necessarily unstressed. Therefore, the object pronoun *het* cannot undergo inversion. Secondly, embedded clauses do not allow inversion, but in embedded clefts, *het* again appears in subject position if the focus is a NP (4a) and in object position if the focus is a pronoun (4b).

(4) a. omdat het jouw kinderen zijn, die huilen
because it your children are, that cry
because it is your children, that cry

b. omdat ik het ben, die dat doet
because I it am, who that does
because it is me who does that

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 the embedded clause is a cleft, the raised subject is *het* (5) (recall that embedded clauses in Dutch are SOV and do not allow for subject object inversion).

[3] In the following, whatever I say about the agreement features of *het* (it) also applies to the pronouns *dat* (that) and *dit* (this).

(5) omdat het jouw kinderen lijken te zijn, die huilen.
because it your children seem(pl) to be that cry
because it seems to be your children, that cry.

Now that we have established that the pronoun *het* is the subject, how can we account for the agreement features of the verb? Example (5) illustrates that both the raising verb and the embedded copula show plural agreement. Following the strict subject verb agreement in Dutch, we have to conclude that *het* is plural in the examples (1a), (4a) and (5).

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 (6-7) and resumptive pronouns (8) (example from (Rullman and Zwart, 1996)). A classic discussion in Dutch linguistics deals with the question which of the constituents is subject in sentences like (6a) (Merckens, 1961; Bos, 1961), where the word order suggests that *dat* is the subject, but subject verb agreement suggests that *schurken* (crooks) 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.

(6) a. Het/dat/dit zijn schurken.
it/that/this are crooks
They are crooks.

b. Het/dat/dit is een mooie vrouw.
it/that/this is a beautiful woman
It/that/this is a beautiful woman.

(7) a. * Het/dat/dit zijn mooi.
it/that/this are beautiful

b. Wat vind je van dit boek? Het is mooi.
what think you of this book(sg.neut) it are beautiful
What do you think of this book? It is beautiful.

c. * Wat vind je van deze auto? Het is mooi.
what think you of this car(comm) it is beautiful

(8) a. Jan en Piet, dat zijn soldaten.
John and Pete that are soldiers
John and Pete, they are soldiers.

het:		
	PRED	'PRO'
	PERS	3
	NUM	(sg)
	GEN	(N)
	PRONTYPE	'cop'

Figure 1: Lexical entry for *het*

b. * Ambtenaren, dat zijn vervelend.
civil servants that are boring

The contrast between the examples (6a) and (7a) and between (8a) and (8b) illustrates that the apparently non-agreeing pronouns can only be used appropriately if an NP object is present. Note that the agreement constraints on the verb have to be constraining equations, not defining equations.

The pronoun itself does not show agreement, but subject verb agreement in example (6a) and resumptive pronoun antecedent agreement in example (8a) indicate that the value for NUMBER on the subject is in fact plural (and GENDER is common). 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 (7b-7c).

This pattern can be accounted for if we assume a lexical entry for *het* as in figure 1,[4] where the pronoun has an optional singular neuter agreement value. This means that if there is a nominal object, agreement between the subject and the object of the copula can override the default agreement value of the pronouns, which is what happens in the plural examples. Since the predicative adjective does not have agreement features, there is no agreement in number or gender between the subject and the adjective: the subject cannot 'get' agreement values from the predicate. As a result, the subject can only satisfy the agreement constraints on the verb by instantiating the default value for number and person: singular neuter. This explains why the examples (7c) and (8b) are out. In addition, the pronoun is has a feature PRONTYPE with value 'cop' (copular). This feature value pair sets apart *het* (it), *dat* (that) and *dit* (this) 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).

[4]The parentheses around the optional features translate to the following disjunction: $(\uparrow \text{NUM}) \vee (\uparrow \text{NUM})=\text{sg}$ and $(\uparrow \text{GEN}) \vee (\uparrow \text{GEN})=\text{neut}$.

2.2 The relative clause

Clefts with a nominal focus have a final relative clause.[5] The relativizer appears to agree in gender with the 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 focus (1c), as it does in adjoined relative clauses (9).[6]

(9) Ik, die jou altijd steun(1sg), zal(1sg) je ook nu helpen
I who you always support(1sg) will you also now help(1sg)
I, who always support you, will also help you now

Alternatively, one could assume that the object, rather than the focus, is the antecedent. This would entail that in the pronoun cases, the antecedent is *het*. The third person agreement on the embedded verb in (1c) would thereby be explained.

Example (10) appears to be a counterexample to this analysis: the embedded verb is plural, whereas the antecedent is *het*. Similarly, the relativizer in (11) is of common gender, while the antecedent is *het*. However, with the lexical entry proposed in figure 1, these examples are no longer problematic.

(10) Wij zijn het, die dat doen
we are it that do that
It is us who do that

(11) Zij is het die hem irriteert
she(comm) is it that(comm) him irritates
She is the one that irritates him

One disadvantage of the object antecedent approach remains: since the focus is sometimes the subject and sometimes the object of the cleft sentence, the discourse function of the relative clause would vary under this analysis: the clause would be

[5] I 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

[6] In old-Dutch and in some bible texts one can also find first person verbs in it-clefts. I do not account for these archaic examples here.

(1) 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

part of the TOPIC if the focus is a pronoun (and subject), and part of the FOCUS otherwise. 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 pseudoclefts 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 (1a) 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 refering to animate versus inanimate objects.[7] Furthermore, free relatives are always singular, with a universal or exhaustive reading, whereas the final clause of a cleft can be plural (1a).

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, we cannot treat the pronoun *het* in clefts as the expletive pronoun. This means that *het* has a PRED feature, which would clash with the PRED 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 his 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 account for 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 (1a) should be plural. With the lexical entry for *het* presented in figure 1, 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.

[7] But note that the reference grammar of Dutch (Haeseryn and others, 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 and others, 1882-1998), column 2517-2518)

That finally brings us to the analysis I propose in this paper: 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 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 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). The next section will show how this analysis can be formalized.

2.3 Formalization

I 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 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. 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.

Like in regular relative clauses, the relative pronoun in the final clause can be embedded (12). These examples are automatically accounted for by the regular relative clause rules.

(12) Het is Jan aan wie ik denk
it is John on who I think
It is John who I am thinking of

The 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). An example c-structure is given in fig. 3, the corresponding f-structure in fig. 4.

The c-structure rules in figure 2 show that the transitive it-cleft is a construction, with construction specific features: the relative clause does not form a $\bar{\mathrm{N}}$ with its

IP	⇒	DP	I’	(CP)
		(↑SUBJ)=↓	↑=↓	(↓TYPE)=$_c$rel
				(↑FOCUS)
				(↑TOPIC ADJ)=↓
				(↑TOPIC PRONTYPE)=’cop’
I’	⇒	I	VP	
		↑=↓	↑=↓	
VP	⇒	DP	V	
		(↑OBJ)=↓	↑=↓	

Figure 2: C-structure rules for nominal clefts

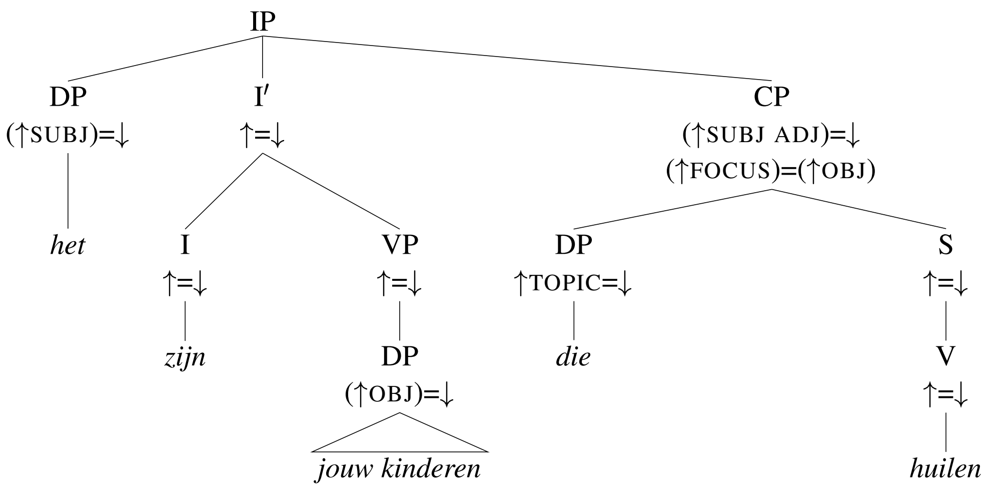

Figure 3: C-structure for (1a) *Het zijn jouw kinderen die huilen*

$$
\left[\begin{array}{ll}
\textsc{pred} & \text{'be-equal-to}\langle(\uparrow\textsc{subj})(\uparrow\textsc{obj1})\rangle\text{'} \\
\textsc{subj} & \left[\begin{array}{ll}
\textsc{pred} & \text{`pro'} \\
\textsc{agr} & \text{3pl} \\
\textsc{adj} & \left\{ \left[\begin{array}{ll}
\textsc{pred} & \text{`cry}\langle(\uparrow\textsc{subj})\rangle\text{'} \\
\textsc{type} & \text{`rel'} \\
\textsc{relpro} & \left[\begin{array}{ll} \textsc{pred} & \text{'pro'} \\ \textsc{agr} & \text{3pl} \end{array}\right] \\
\textsc{subj} & [\;]
\end{array}\right] \right.
\end{array}\right] \\
\textsc{obj} & \left[\begin{array}{ll}
\textsc{pred} & \text{'child-of}\langle(\uparrow\textsc{poss})\rangle\text{'} \\
\textsc{agr} & \text{3pl} \\
\textsc{poss} & \left[\begin{array}{ll} \textsc{pred} & \text{'pro'} \\ \textsc{agr} & \text{2sg} \end{array}\right]
\end{array}\right] \\
\textsc{focus} & [\;] \\
\textsc{topic} & [\;]
\end{array}\right]
$$

Figure 4: F-structure for (1a) *Het zijn jouw kinderen die huilen*

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 expanded them to also cover cleft sentences. The only component that was added is the optional relative clause with the construction specific information.

The analysis leaves the fixed 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.

Agreement in number and gender between the two arguments of the copula appears to be a general but violable (13) principle. I assume some form of optimization here. Alternatively, one could formulate the following construction specific constraint on the CP in the c-structure rules: (↑SUBJ NUM)=(↑OBJ NUM) (and a similar one for GEN). However, these would not account for gender agreement in other copular sentences such as (14).

(13) Als zij mij waren, dan zouden ze hetzelfde doen
if they me were than would they the same do
If they were me they would do the same

(14) a. Dit boek is het enige dat ik gelezen heb
this(neut) book is the(neut) only that(neut) I read have
This book is the only one that I read

b. Deze rok is de enige die me past
this(comm) skirt is the(comm) only that(comm) me fits
This skirt is the only one that fits me

The rules in figure 2 do not specify the verb to be copular. This can be achieved by constraining the complement to be of the copular-specific type. Because I did not discuss the status of the complement of the copula earlier and simply called it OBJ, I will do the same here. Alternatively, it could be specified on the final clause as a construction specific constraint.

I do not account for the distribution of the two argument structures of the transitive cleft in this paper. The rules in figure 2 generate both argument structures for both pronouns and semantic NPs. It is assumed that general constraints penalize pronouns—and in particular focused pronouns—in object position, excluding sentences like (15). This assumption is supported by the fact that the same effects can be observed in non-cleft copular sentences. In (Haeseryn and others, 1997) several non-cleft copular constructions are mentioned with their particular usage of pronouns and from this list we can conclude that pronouns are generally avoided in non-subject positions. If *both* arguments are pronominal, then the local or the stressed pronoun is in subject position. An account for one example of this phenomenon is sketched in (Coppen, 1996), but it seems possible to formulate a more general account for the distribution of pronouns in copular sentences based on a pronominal hierarchy like the one (Müller, 2002) proposes for German.

(15) 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

The analysis of relative clause clefts as instances of transitive copular sentences is supported by the fact that this type of cleft can be rephrased as a canonical specificational sentence, see example (1a) and the rephrase (16).

(16) Diegenen die huilen, zijn jouw kinderen
the-ones who cry are your children
Those who cry are your children

3 Intransitive clefts

So far, we have looked at it-clefts with nominals in the focus position. But not only noun phrases can be clefted: also PPs (1d), CPs (1e), AdvPs (17a) and arguably APs (17b) can appear in it-clefts.

(17) a. Het was toen pas, dat ik een vermoeden kreeg
it was then only, that I a suspicion had
It was not until then, that I became suspicious

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

3.1 Differences between transitive and intransitive clefts

The structure of these clefts is different from the it-clefts with a nominal focus. In the first place, they have a final complementizer clause instead a relative clause. This clause is always headed by the complementizer *dat* (that). That-clauses, but not relative clauses, can function independently as an argument in Dutch.

Secondly, while the transitive cleft had a variable argument structure, this second type of cleft has only one possible word order and argument structure: the pronoun *het* (it) is always in subject position. Also, the agreement on the copula is invariantly third person singular. In contrast to the transitive cleft, this construction does not allow the pronoun *het* to be replaced by a demonstrative.

Finally, in clefts with a complementizer clause instead of a relative clause the focus may fill one of the argument functions of the embedded verb, as in example (18).

(18) Het is aan hem dat ze denkt
het is of him that she thinks
It's of him that she thinks

Because the focus is not an NP, it cannot be the OBJ of the copula. Even if we allowed as object all categories that can serve as the complement of a copula in predicative sentences, there are still examples that fall out, e.g. sentence (18) above.

An analysis along the lines of the nominal clefts would not be appropriate either, because the meaning is different. In contrast to the relative clause clefts, the complementizer clefts cannot be rephrased as canonical specificational sentences (19b), even if we transformed the that-clause into a locational free relative (19c). The best rephrase would be the simplex sentence in (19d).

(19) a. Het was in Amsterdam, dat ik hem voor het eerst ontmoette.
it was in Amsterdam that I him for the first met
It was in Amsterdam that I first met him.

b. * Dat ik hem voor het eerst ontmoette was in Amsterdam
that I him for the first met was in Amsterdam

c. * Waar ik hem voor het eerst ontmoette was in Amsterdam
where I him for the first met was in Amsterdam

d. Ik ontmoette hem voor het eerst in Amsterdam
I met him for the first in Amsterdam
I first met him in Amsterdam

So we need two distinct analyses for the two types of clefts. A similar claim has been made for English by Pinkham and Hankamer (1975). They account for the differences between nominal and non-nominal clefts in English on the basis of the dual derivation principle. This principle states that nominal clefts can be base generated, whereas non-nominal clefts have to be derived from the clause. This idea has been incorporated in more recent cleft analyses e.g. (Merchant, 1998). In contrast to the dual derivation analysis, the account presented here is non-derivational. The differences between the two types of clefts are argued to be the result of the difference in argument structure of the copula: in section 2 I argued that the relative clause cleft is a construction with the transitive (specificational) copula. In this section I will proceed to argue that the complementizer clause cleft is not based on the transitive copula, but on the intransitive (existential) copula.

3.2 The intransitive analysis

The complementizer cleft consists of the copula and three constituents: the pronoun *het*, the focus and the final clause. As noted before, the subject *het* cannot be replaced by a demonstrative or any other NP in complementizer clefts, which is the most important test for expletive subjects. It also cannot be stressed and does not support emphatic reflexives, additional tests for expletives (Postal and Pullum, 1988). Conclusion: *het* is an expletive pronoun.

The second constituent is the focus. It cannot be analyzed as the object of the copula, because it is not an NP. But if it is not OBJ, then what is it? In examples like (18), the focus is an argument of the embedded verb, which shows that the constituent is extracted from the that-clause. Extraction can also account for the other examples by mean of adjunct extraction instead of argument extraction. In other words: the focus in itself is not an argument of the matrix verb, but an extracted constituent of the complementizer clause.

The observation that the focus can be associated with an argument of the embedded verb has inspired previous analyses of the cleft construction. Transformationalists derived complementizer clefts (19a) from a canonical sentence (19d) out of which the clefted element was moved (Pinkham and Hankamer, 1975; Emonds, 1976). Pollard and Sag (1994) account for this fact in a non-transformational way. They assume a special lexical entry for *be* for clefts with on the subcategorization list *het*, an XP and a complementizer clause with that XP on slash.

The third and last constituent is the complementizer clause. It 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. This is also illustrated in (20). The example is similar to sentence (1d), but now without extraction of the focus out of the clause, so that it is a regular that-clause object. It is not only syntactically marked, but the meaning is also different from the meaning of the cleft sentence, because the pronoun is interpreted referentially instead of expletive.

(20) ? Het was dat ik hem in Amsterdam voor het eerst ontmoette
it was that I him in Amsterdam for the first met

Alternatively, I 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 focus) is 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 can be seen in sentences with CPs in canonical subject position (21).

(21) Dat we gewonnen hebben is nog niet zeker
that we have won 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. 5. The corresponding f-structure is in fig. 6.

3.3 Formalization

The c-structure rules for intransitive clefts are given in figure 7. The rules also account for sentences with embedded 'gaps' in the complementizer phrase (22).

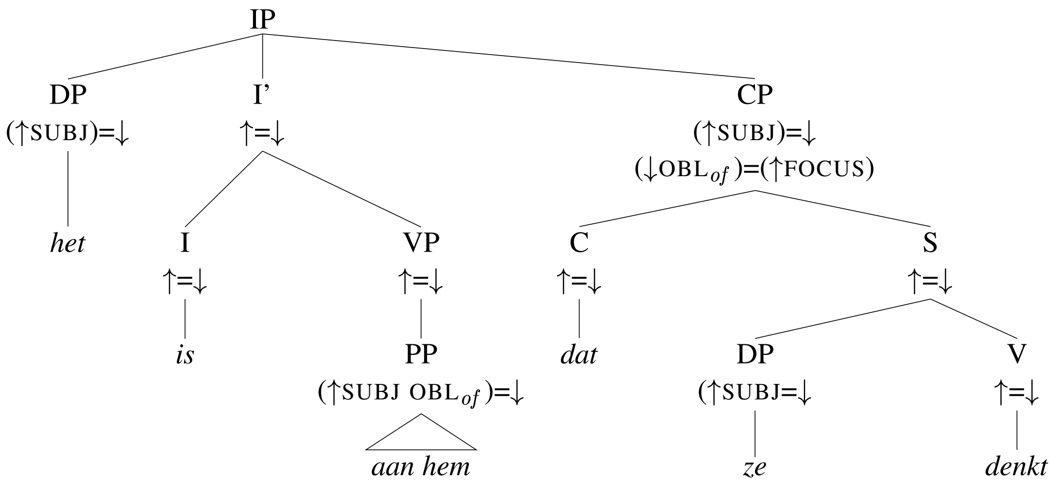

Figure 5: c-structure for 18 *Het is aan hem dat ze denkt*

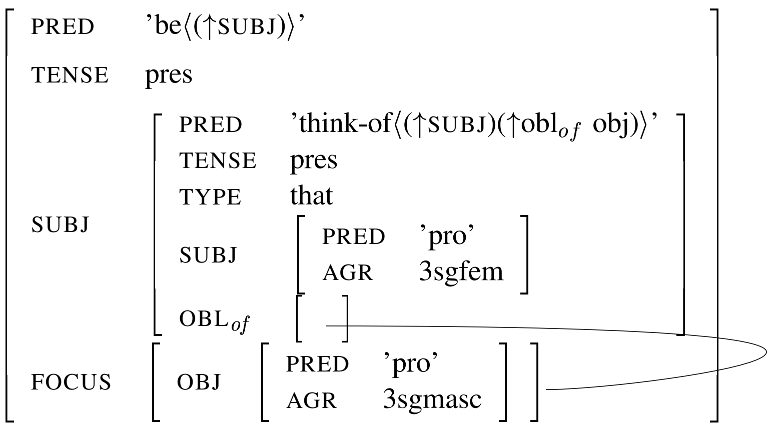

Figure 6: f-structure for 18 *Het is aan hem dat ze denkt*

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 (23) and thus separates the two phrases.

(22) Het was in Amsterdam dat ze hoopte hem te kunnen ontmoeten
it was in Amsterdam that she hoped him to can meet
It was in Amsterdam that she hoped she could meet him

(23) Het moet in Amsterdam geweest zijn dat ze hem ontmoette
it must in Amsterdam been be that she met him
It must have been in Amsterdam that she met him

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. I did not specify the argument function of the focus either, which means that the complementizer clause has to be instantiated to determine the syntactic function of the focus (or the coherence principle is violated).

In section 3.1 I stated that the closest rephrase of the intransitive cleft (19a) is the simplex sentence (19d). 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).

Note that the intransitive analysis would be inappropriate for the it-clefts with a relative clause, which I analyzed as transitive copular sentences, because the intransitive analysis crucially depends on the subject being expletive and I 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 rephrases. The two have in common that given information is extraposed to focus new information.

4 Conclusion

I have accounted for the syntactic differences between the Dutch it-cleft with a nominal focus and those with a non-nominal focus by analyzing the first as an

$$
\begin{array}{lllll}
\text{IP} & \Rightarrow & \text{DP} & \text{I'} & \text{(CP)} \\
 & & (\uparrow\textsc{subj}){=}\downarrow & \uparrow{=}\downarrow & (\downarrow\textsc{type}){=}\text{that} \\
 & & & & (\uparrow\textsc{subj}){=}\downarrow \\
 & & & & (\downarrow\text{CP}\textsc{ath}){=}(\uparrow\textsc{focus}) \\
\\
\text{I'} & \Rightarrow & \text{I} & \text{VP} & \\
 & & \uparrow{=}\downarrow & \uparrow{=}\downarrow & \\
\\
\text{VP} & \Rightarrow & \text{(CleftP)} & \text{(V)} & \\
 & & (\uparrow\textsc{focus}){=}\downarrow & \uparrow{=}\downarrow &
\end{array}
$$

$\text{CP}\textsc{ath}{=}\textsc{xcomp}{*}\{\textsc{obl}_{\theta}|\textsc{adj}\}$
$\text{CleftP}{=}\{\text{AP,PP,CP}\}$

Figure 7: C-structure rules for non-nominal clefts

instance of the transitive copula and the second as an instance of the intransitive copula.

I 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, I also accounted for the subject verb agreement pattern in this type of cleft. In addition, I 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.

In the second type of cleft, the intransitive cleft, all phrases are associated with the subject function of the copula: the 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.

I did not have to stipulate construction specific lexical entries for the copula, because I expanded the existing rules for the transitive and intransitive uses of the copula with optional construction specific elements. Both types of clefts involve extraposition of given information in order to focus new information.

Two questions were left unanswered. In the first place, I 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 the second place, I did not account for the fact that the two argument structures that are found in transitive clefts are in complementary distribution: one is used if the focus is pronominal and the other if the focus is a full NP. I leave these questions open for future research.

References

Akmajian, Adrian. 1970. On Deriving Cleft Sentences from Pseudo-Cleft Sentences. *Linguistic Inquiry*, 1(2):149–169.

van der Beek, Leonoor. 2001. It-clefts in Dutch. Presentation at ADL01 in Paris.

Berman, Judith. 2001. On the cooccurrence of es with a finite clause in german. In T. Kis and D. Meurers, editors, *Constraint-Based Approaches to Germanic Syntax*. CSLI, Stanford, CA.

Bos, G. 1961. Dat zijn kooplieden. *De Nieuwe Taalgids*, 54:23–27.

Coppen, P.A. 1996. Als dit 't nou eens was. Verschenen in NEDER-L, sept 1996. `http://www.neder-l.nl/archieven/miniatuurtjes`.

Dalrymple, Mary. 2001. *Lexical Functional Grammar*. Academic Press.

Declerck, Renaat. 1988. *Studies on copular sentences, clefts and pseudoclefts*. Leuven University Press/Foris Publications.

Emonds, J.E. 1976. *A Transformational Approach to English Syntax*. Academic Press, New York.

Falk, Yehuda. 2001. *Lexical-functional grammar: an introduction to parallel contraint-based syntax*. CSLI Publications.

Haeseryn, W. et al., editors. 1997. *Algemene Nederlandse Spraakkunst*. Nijhoff, Groningen.

Higgins, F.R. 1976. *The pseudo-cleft construction in English*. Ph.D. thesis, Massachusetts Institute of Technology.

Jespersen, O. 1927. *A Moderm English Grammar III*. George Allen and Unwin.

Merchant, Jason. 1998. 'Pseudosluicing': Elliptical clefts in japanese and english. In A. Alexiadou et al., editors, *ZAS Working Papers in Linguistics*. Zentrum für Allgemeine Sprachwissenschaft: Berlin.

Merckens, P.J. 1961. Zijn dat kooplieden of zijn kooplieden dat? *De Nieuwe Taalgids*, 54:152–154.

Müller, G. 2002. Harmonic alignment and the hierarchy of pronouns in german. In H. J. Simon and H. Wiese, editors, *Pronouns - Grammar and Representation*. John Benjamins, pages 205–232.

Pinkham, Jessie and Jorge Hankamer. 1975. Deep and shallow clefts. In R.E. Grossman, L.J. San, and T.J. Vance, editors, *Papers from the Eleventh Regional Meeting of the Chicago Linguistic Society*, pages 429–450.

Pollard, Carl and Ivan Sag. 1994. *Head-driven Phrase Structure Grammar*. University of Chigago / CSLI.

Postal, Paul M. and Geoffrey K. Pullum. 1988. Expletive NPs in subcategorized positions. *Linguistic Inquiry*, 19:635–670.

Rizzi, Luigi. 1997. The fine structure of the left periphery. In L. Haegeman, editor, *Elements of Grammar: Handbook in Generative Syntax*. Kluwer Academic Publishers, Dordrecht.

de Rooy, J. 1970. Onzijdige voornaamwoorden en het naamwoordelijk gezegde. *De nieuwe taalgids*, 63:181–186.

Rullman, Hotze and Jan-Wouter Zwart. 1996. On saying *Dat*. In Roel jonkers, Edith Kaan, and Anko Wiegel, editors, *Language and Cognition 5*, Groningen.

Smits, R.J.C. 1989. *Eurogrammar; The Relative and Cleft Constructions of the Germanic and Romanic Languages*. Foris, Dordrecht.

de Vries, M. et al., editors. 1882-1998. *Woordenboek der Nederlandsche taal*. Nijhoff [etc.].