# On D-pronouns and the Movement of Topic Features

## Eric Hoekstra

## 1. Introduction[1]

This explorative article deals with the distribution of topic pronouns in Dutch. It is argued that the behaviour of such pronouns is not predicted by theories incorporating the generalisation that a pronoun is free in the clause in which it occurs. The topic pronoun must instead be free in the matrix sentence. We derive these facts from the assumption that the topic pronoun is moved to the beginning of the matrix sentence in LF. We then go on to characterise the set of positions in which antecedents for topic pronouns are found. These positions, it is suggested, function as escape hatches for topic features, in which they can access the larger structure in which sentences are integrated.

## 2. Where topic pronouns differ from ordinary pronouns

### 2.1 Basics

A typical marker of spoken language in Dutch and Dutch dialects is the topic pronoun *dat* or *die* 'that'. This pronoun tends to be absent in written language whereas it is abundantly present in spoken language. Consider for example the following sentences (where a topic pronoun is glossed as "TP"):

(1)  a.  De minister wist  er  van
           the minister knew it  about
           "The minister knew about it."
     b.  De minister die  wist  er  van
           the minister TP  knew it  about
           "The minister knew about it."

These two sentences differ minimally in that the subject *de minister* is doubled by the topic pronoun *die* in the (b)-sentence but not in the (a)-sentence. The

topic pronoun *die* is logically superfluous but that does not mean it does not have a grammatical function (De Vries 1910-1912). The (b)-sentence would not be used in written language, nor in formal spoken language. The (a)-sentence, on the other hand, is used in either spoken or written language. I will refer to the constituent doubled by the topic pronoun as a dislocated constituent. Notice that there is normally not a prosodic break between the dislocated constituent and the topic pronoun.

The morphological paradigm of topic pronouns is identical to that of demonstrative pronouns used for non-proximate reference. It is quite simple, as given below:

(2)  SINGULAR                        PLURAL
     Non–neuter      Neuter
     *die* 'that'     *dat* 'that'      *die* 'those'

As demonstratives, *die* is used with nouns taking the definite article *de* (non-neuter nouns) and *dat* is used with nouns taking the definite article *het* (neuter nouns).[2] Non-neuter and plural have fallen together, not only in this case, but also in the case of proximate demonstrative pronouns, definite articles and adjective inflection. It must also be mentioned that the neuter demonstrative pronoun *dat* is homophonous to the complementiser ("that"), as in English.

There are also topic pronouns for time and place adverbials and for conditional clauses:

(3)  a.  Morgen    dan      kom ik bij je   langs
         tomorrow  TP:then  come I   at  you  along
         "Tomorrow, I'll visit you."
     b.  Gisteren   toen     kwam ik bij je   langs
         yesterday  TP:then  come  I   at  you  along
         "Yesterday, I'll visit you."
     c.  In Den Haag   daar    wil      ik graag  wonen
         in  The Hague  TP:there would  I   indeed  live
         "In The Hague, I would like to live."
     d.  Als 1  en   1 3 is     dan      heb je   misschien gelijk
         if   1  and  1 3 equals TP:then  have you  maybe      right
         "If 1 and 1 equal 3 then you could be right."

The topic pronoun for conditional clauses is identical to the topic pronoun for time adverbials. Topic pronouns all begin with a D- in Dutch, with the exception of the past tense topic pronoun *toen* 'then'.[3] I will in this study be mostly concerned with topic pronouns for DP. Both the term topic pronoun and the term D-pronoun are used in this study.

## 2.2 *Surprises*

The behaviour of topic pronouns is not predicted by condition B of the Binding theory nor by principles designed to replace it, such as Reinhart's (1981, 1983)

pragmatic principle. The reason is simple: most theories are designed to capture the behaviour of "normal" pronouns. Normal pronouns obey the generalisation that they are free in the clause in which they occur, and may be bound outside. Topic pronouns simply behave different, and thus the behaviour of topic pronouns is not predicted by any theory incorporating the generalisation that pronouns are free in the clause in which they occur. We will now go on to present some facts, in which topic pronouns are contrasted with normal pronouns. To begin, consider the following (judgments reflecting spoken language, not written language where D-pronouns are taboo):

(4)   a. Marie   zei   tegen Piet$_i$ [dat  ik  hem$_i$/*die$_i$  niet  mag]
        Marie   told  to     Piet   that I   him/TP      not   like
        "Marie told Piet that I did not like him."

    b. Marie$_i$   zei  tegen Piet$_i$ [dat . ik haar$_i$/*die$_i$  niet  mag]
        Marie     told  to     Piet   that I   her/TP      not   like
        "Marie told Piet that I did not like her."

    c. Jan$_i$  vond  het huis  [dat ze    hem$_i$/*die$_i$ wilden   verkopen]niet mooi
        Jan  found  the house that they him/TP     wanted sell       not  nice
        "Jan did not like the house which they wanted to sell him."

    d. *D-pronoun domain generalisation*
        A D-pronoun is free in the domain of the matrix sentence in which it occurs (not just the clause)

The domain in which pronouns must be free according to condition B and other theories incorporating the clause generalisation is bracketed in the examples above. The observation is that D-pronouns must be free is a much larger domain than just the clause in which they occur. In the (c) sentence the D-pronoun is separated from its antecedent by two clause boundaries and a DP boundary, yet coreference is excluded. The tentative generalisation is given in (d).

The notion "free" is normally defined in terms of the c-command:

(5)   A is free from B iff B does not c-command A,
    where A is coindexed with B.

The examples above are problematic for the definition of the syntactic domain in which condition B applies but they are not problematic for the application of the notion c-command to these examples.

## 2.3 Proposal

To account for these facts, we can stipulate a different domain definition for topic pronouns. This covers the facts, but it does not buy us anything else. Instead, I would like to propose a change in the conceptual definition of condition B, which automatically predicts different binding domains for topic pronouns and ordinary pronouns.

The idea is based on the minimalist insight (Chomsky 1995) that if a process is overt is one language (feature F is strong and therefore it triggers movement before PF), it must be covert in another (feature F is weak and therefore it does not trigger movement until LF). After all, this is the essence of the Minimalist program. Now, pronouns are moved in PF in some languages into the functional domain as clitics. If they are not moved in PF they must be moved in LF. Therefore we claim, by theory-internal reasoning, that the following holds:

(6) *Pronoun Movement Hypothesis* (follows from Minimalist Program)
   All languages create pronominal clusters in the functional domain in LF.

What is important is that pronouns, under this proposal, must move outside the domain of thematic position. This will derive a configuration where even the pronoun coindexed with the lowest (in terms of c-command) theta-role crosses the theta position of the highest pronoun:

(7) Pronominal cluster (P:1 ... P:n) ... Th-pos:1 ... Th-pos:n
   where 1 ... n is used to represent binding,
   where the pronominal cluster c-commands Th-pos:1,
   Th-pos:1 c-commands Th-pos:2, and so on

We thus assume that a clitic cluster is in LF always higher (in terms of c-command) than the domain in which a verb discharges its theta roles: this is a feature of most, if not all, analysis of clitic clusters.

Every pronoun in the cluster in (7) must be able to bind its trace in theta position. The consequences of this are given in (8):

(8) *Consequence of the Pronoun Movement Hypothesis*
   If a pronoun is coreferential with another argument in the domain of its chain, that pronoun will bind two th-positions, thus have two theta roles.

(9) *Theta Criterion:* a chain may not have two or more th-roles

(10) *Consequence of Pronoun Movement and Theta Criterion*
   A pronoun may not be coreferential with another argument in the domain defined by its movement chain

A relevant instantiation of (10) is that a pronoun may not be coreferential with another argument of the same verb (the function of reflexives is to circumvent this, for example in the manner proposed by Reinhart & Reuland (1991)).

We have now derived condition B from the minimalist hypothesis that pronoun clusters are created in LF on the analogy of clitic clusters in PF. Notice that we do not have to formulate condition B as an independent principle. The clause-boundedness of condition B comes for free as a result of the clause-boundedness of pronoun movement in LF.

We are now in a position to develop a conceptual analysis of the D-pronoun facts. This is given in (11):

(11) *Fact:* D-pronouns must be free in a larger domain than ordinary pronouns, namely in the domain of the matrix sentence

(12) *Consequence of (10) and (11)*
D-pronouns undergo long-distance LF-movement to the beginning of the matrix sentence

Thus we have proposed an account capturing both the binding facts concerning ordinary pronouns and the binding facts concerning D-pronouns in a non-trivial manner, i.e. we do not stipulate the domain in which a pronoun must be free. The size and location of the domain is defined by the chain representing the movement which the pronouns undergo in LF.

## 3. Who moves in left-dislocation constructions?

### 3.1 Is the dislocated phrase moved or the topic pronoun?

Consider a simple case of left-dislocation with topicalisation of the D-pronoun:

(13)  Jan  die  ken   ik  niet
      Jan  TP   know  I   not

It might also be suggested (Grohmann 1997) that the dislocated phrase *Jan* is moved, and that the topic pronoun represents the spelling out of a trace. I will refer to this analysis as the spell-out analysis. I will argue first against a conventional spell-out analysis. On the basis of the arguments presented, I will develop a refined version of the analysis, incorporating insights from the spell-out analysis but avoiding the problems.

### 3.2 Arguments against movement of the dislocated phrase

If the dislocated phrase is what is moved we expect the dislocated phrase to satisfy selection requirements in the base position of the chain. If the topic pronoun is what is moved then we expect the topic pronoun to satisfy selection requirements in the base position of the chain. The two options are schematised below:

(14)  a.   Dislocated $XP_i$   [ TP-pronoun ... [ ... $t_i$ ... ]
      b.   Dislocated XP    [ TP-pronoun$_i$ ... [ ... $t_i$ ... ]

Of course, the element that moves will have to satisfy the selection criteria downstairs. Thus the analysis in (a) predicts that the dislocated XP will satisfy selection criteria downstairs. The (b) analysis predicts that the topic pronoun satisfies selection criteria downstairs. We will now present three cases indicating that the (b) analysis is correct, not the (a) analysis (from Hoekstra & Zwart 1994).

Case 1. Consider the following contrast:

(15) a. *Ik   doe   niet   [boeken   lezen]
       I     do    not    books     read
       "I don't read books."

   b. [Boeken   lezen]   dat   doe   ik   niet
       books     read     TP   do   I    not
       "Read books, I don't."

   c. Ik   doe   dat   niet
       I    do    TP   not
       "I don't do that."

The infinitival IP-constituent *boeken lezen* 'read books' may not occur as a complement to *doen* 'do', as the (a)-sentence shows, because this verb does not categorially select infinitival IP. If *boeken lezen* 'read books' occurs as a left-dislocated constituent at the periphery of the clause, however, and is doubled by a topic pronoun, the structure is fine. These facts are explained if it is the topic pronoun which moves, after it has satisfied the selection criteria downstairs. The dislocated phrase cannot have been moved because it violates the selection criteria downstairs, as indicated in (a).

Case 2. A comparable example comes from Frisian. The verb *wenje* 'to live (somewhere)' requires a PP complement, but a dislocated phrase may have the form of a DP (Hoekstra 1995):

(16) a. Reduzem  dêr       woe   ik wol wenje
       Reduzem  TP:there would I   sure live
       "In Reduzem, I sure would like to live."

   b. *Ik woe   wol  Reduzem  wenje
       I   would  sure Reduzem  live
       "I sure would like to live in Reduzem."

   c. Ik woe   wol dêr   wenje
       I   would  sure there live
       "I sure would like to live there."

If the dislocated phrase had been moved, it would violate the selection criteria downstairs, as shown by (b), and the sentence (a) would be ungrammatical. If the locative D-pronoun is what moves, it satisfies the selection criteria downstairs, as shown by (c), and (a) is grammatical, as expected.

It could be argued that these examples involve so-called "hanging topic" (HT) constructions. HT constructions in German (Grohmann 1997) have the following properties (in the gloss REFL stands for "reflexive"):

(17) a. The dislocated XP cannot be bound into.
   b. The dislocated XP does not Case-agree with the topic-pronoun.

(18) a. *[Der Wagen von sich] den hat er verkauft
     the car   of  REFL TP has he sold
     "His car, he sold."

   b. [Den Wagen von sich] den hat er verkauft
     the car   of  REFL TP has he sold
     "His car, he sold."

The (a) sentence shows that a HT-construction cannot be bound into and does not case-agree with the topic pronoun. The (b) sentence illustrates a non-HT construction, according to Grohmann. In Dutch, case is not visible on DP but we can apply the binding test to see whether the examples discussed by us are HT-constructions.

    Consider the following sentences, in which the dislocated element violates selection criteria, yet binding into it is allowed:

(19) a. [Elkaar$_j$   helpen]$_i$ dat doen ze$_j$  hier niet t$_i$
     each other help   TP do   they here not
     "Help each other, they don't do that here."

   b. [Syn$_j$ berteplak] dêr$_i$  soe   elkenien$_j$  wol t$_i$ wenje wolle
     his   birth place there would everybody sure  live   want
     "Everybody would sure like to live in the place of his birth."

(20) a. *Ze   doen hier niet elkaar    helpen
     they do   here not each other help
     "They don't help each other here."

   b. *Elkenien   soe   wol syn berteplak  wenje wolle
     everybody would sure his birth place live   like
     "Everybody sure would like to live in the place of his birth."

(21) a. Ze   doen dat niet
     they do   that not
     "They don't do that."

   b. Elkenien   soe   wol dêr    wenje wolle
     everybody would sure TP:there live   like
     "Everybody sure would like to live there."

Binding is possible; hence the counterexamples presented here cannot be explained away as involving HT-constructions.

### 3.3 The paradox

We now have an apparent paradox: the facts we presented indicated that the topic pronoun satisfies the selection criteria downstairs. Nevertheless, the dislocated constituent may be bound into. Although an account of reconstruction and binding is beyond the scope of this paper, we would like to point out that Grohmann's assumption that binding requires reconstruction to a simple c-command configuration might be mistaken. To illustrate, consider the following:

(21) c. [What they$_i$ did] was eat each other's$_i$ sandwiches

    d. [Every farmer who owns a donkey$_i$] loves it$_i$

In both cases there is no way in which reconstruction can apply so as to bring the antecedent in a c-commanding position. The examples make clear that the binding mechanism may work in intricate ways. A full treatment of binding and reconstruction falls outside the scope of this paper.

    Earlier we argued that pronouns must be free in the domain of their chains, and that topic pronouns move to the beginning of the matrix sentence in LF. However, we also observe that a topic pronoun can coindex with a left-dislocated constituent:

(22) a. Jan die ken   ik niet

       Jan TP know I  not

       "Jan I don't know."

    b. Ken  je   Jan? Die ken  ik niet

       know you Jan TP know I  not

       "Do you know Jan? I don't know him."

If our account is correct, then we are bound to conclude that the dislocated phrase in the (a) sentence is outside the domain defined by the chain of the topic pronoun (as is evident in the case in the (b) example). This means that a movement analysis for the dislocated element itself is out of the question. We arrived at the same conclusion on the basis of the facts concerning selection criteria.

    Coindexation between the left-dislocated element and the topic pronoun is not a relation between members of a chain, but it is a relation between two chains. This is immediately clear from the (b) sentence. If we want to generalise over both cases, we must conclude that the (a) sentence also involves two chains.

    The relation between the two chains could either be coreference (cf. Heim 1982) or chain composition (the latter operates in parasitic gaps). A test for coreference since Reinhart (1983 and elsewhere) is that strong quantifiers cannot corefer. This test is applied to German data by Grohmann (1997:15), and the same facts are valid for Dutch:

(23) a. *Niemanden den hat   sie geküßt (German)

       nobody     TP has  she kissed

    b. *Niemand   die heeft ze  gekust (Dutch)

       nobody     TP has  she kissed

(24) a. Sie hat   niemanden geküßt     (German)

       she has   nobody   kissed

    b. Ze  heeft niemand   gekust     (Dutch)

       she has  nobody   kissed

The failure of D-pronouns to corefer with strong quantifiers is an indication that coreference, in some sense, is involved, a weaker concept than binding.

Chain composition, on the other hand, is what relates a parasitic chain to a non-parasitic chain (Chomsky 1982). The defining characteristic is that the parasitic chain is A-free from the chain on which it parasitises (where "e" stands for the parasitic gap):

(25) a. $Who_i$ did [a friend of $e_i$] please $t_i$
  b. *$Who$ $t_i$ pleased [a friend of $e_i$]

(26) a. $Who_i$ did you phone [without talking to $e_i$]
  b. *$Who$ was phoned up [without talking to $e_i$]

In the (a) examples, the parasitic position, indicated by "e" is A-free from the non-parasitic chain. In the (b)-examples, "e" is c-commanded by the trace in A-position, hence A-bound.

We could consider chain composition to be the formalisation of coreference. We can extend the concept of chain composition so as to apply to the relation between a dislocated phrase and a topic pronoun on the assumption below:

(27) The dislocated phrase is not in an A-position.

If (27) holds, chain composition can apply to express coreference between the dislocated phrase and the topic pronoun.

(27) is not implausible, by the following argument. An A-position is an A-position by virtue of being linked up to a predicate, either by means of case or by means of theta role. A left-dislocated constituent, being base-generated, is not linked up to any predicate. Therefore it is not implausible to suppose that left-dislocated elements can undergo chain composition with the topic pronoun. Below we will suggest that the topic pronoun is related to its antecedent by means of movement of zero topic features.

## 4. Topic pronouns and the positions in which their antecedents are found

### 4.1 Textual structure and the operations "move" and "merge"

It was argued above that topic pronouns must be free in the matrix sentence in which they occur. The question therefore arises:

(28) Where does a topic pronoun find its antecedent?

Normal syntactic principles apply within the clause, or within the sentence but we have not explored the question whether they apply outside the sentence. One thing is clear: it cannot be the case that matrix sentences exist in isolation, that they are not part of a larger, say textual structure.

The theoretical interest of topic pronouns is that they both have a function within the sentence and outside, as a means of connecting up sentences. Thus they raise the issue of how sentences connect up with each other and are integrated into a coherent and structured whole, the textual structure. It cannot be doubted that sentences are fused into larger units. Hoeksema & Napoli (1993)

discuss an interesting construction which can either be expressed as one sentence or as two:

(29) a. The sun was so hot (that) I fainted.
     b. I fainted, the sun was so hot.

In the (a)-sentence, *so* links up two clauses. In the (b)-sentence, *so* links up two sentences. As a result, bound variable binding is possible in the (a)-type sentence but not in the (b)-type, as shown below:

(30) a. Nobody was so hot (that) he fainted.
     b. *Nobody fainted, he was so hot.

The sentence boundary in the (b) sentence blocks bound variable binding. The meaning of (a) is very similar to that of (b). Hence the two sentences in (b) and the two clauses in (a) must ultimately receive very similar semantic representations. Within the Minimalist Program, parts of sentences are integrated in the structural representation of the sentence by means of the operations "merge" and "move". The question, how are sentences integrated into a larger structure, translates into the Minimalist Program as follows: does the creation of textual structure also take place by means of the operations "merge" and "move"? If not, how can we express the relation between (29a) and (29b)? Furthermore, we can ask ourselves the question: which syntactic processes occurring within the sentence can apply between sentences? The answer is none: most processes are clause-bound, and therefore, sentence-bound, given that a sentence-boundary is an even stricter boundary than a clause boundary. Even processes that appears to cross clause-boundaries, like wh-movement or bound variable binding, cannot cross sentence-boundaries. This is shown below:

(31) a. Who$_i$ do you think [that we saw t$_i$]
     b. Nobody$_i$ thinks [that he$_i$ will win]

(32) a. *Who$_i$ was it bad luck. You consider t$_i$ responsible for it?
        "It was bad luck. Who do you consider responsible for it?"
     b. *Nobody$_i$ laughed. He$_i$ did not have any sense of humour.
        "Nobody laughed. They did not have any sense of humour."

The relation between a topic pronoun and its antecedent, coreference or chain composition, is able to cross sentences, as shown by (22b) repeated as (33):

(33) Ken   je    Jan$_i$? Die$_i$ ken    ik niet.
     know  you   Jan     TP    know   I    not
     "Do you know Jan? I don't know him."

The answer to the question of where D-pronouns find their antecedent must therefore be: in a larger structure, call it the textual structure. If coreference plays an important part in this process then coreference is of more interest than we might think. Topic pronouns also play a role in the sentence. Thus topic pronouns might provide insight into the way sentences are integrated into larger

structures. The human mind does not deal with sentences in isolation, it places them in a larger structure. Topic pronouns may thus ultimately provide insight into the way in which the human mind structures text. I will go on to argue that the relation between a topic pronoun and a dislocated XP is not subject to island constraints, because the relation is in part created at the higher structural level of the text, rather than at the level of the sentence.

## 4.2 Island violations

Consider the following German example from Grohmann (1997), used by him to illustrate the claim that left-dislocation violates island constraints:

(34) [Diesen Frosch]$_i$ fragt sich     der Bauer, ob     der König den$_i$ mag
     this     frog     asks   himself the farmer whether the king   TP    likes
     "This frog, the farmer wonders whether the king likes him."

Grohmann assumes that the dislocated element has been moved, and that the topic pronoun was the spell-out of a trace. Hence this constituted an island violation. For us, the island violation is in LF, since in LF the D-pronoun will move to the front of the matrix sentence.

     The problem is clear: topic pronouns cannot be moved to the beginning of the sentence without violating island constraints (both under our analysis and under Grohmann's). The problem is reminiscent of the problem of Wh-movement in the seventies: nearly all syntactic processes (passivisation, agreement) are clause-bound, yet Wh-movement is not. The problem has been solved by means of successive-cyclic movement through licensed escape-hatches. The definition of escape hatches depends on the version of the theory (barriers, licensed position), the insight is the same.

     The problem posed by topics is of the same nature, and I would solve it in a similar manner: by the postulation of an escape hatch for topic features, making it possible for the XP-topic pronoun relation to apparently violate movement constraints. The landing site of the pronominal topic features must be an appropriately licensed position, that is, a topic position. My claim is that the position of the topic "diesen Frosch" is licensed in the sentence above by the bridge verb "fragen". The assumption is that only bridge verbs can license topic positions (like only bridge verbs can license escape hatches for successive-cyclic movement).

     The first argument for this claim is semantic. When you "ask something", you ask something *about something else*. This "about something else" is a topic position licensed as an argument of the verb. Thus it makes sense semantically, to suppose that bridge verbs like "think", "say" and so on have a topic position as part of their array of arguments.

     Second argument. If the topic is part of the argument structure of the matrix verb, then it is at least possible that the matrix verb may impose some formal

selection restrictions. In German, this is not visible, since the topic takes along its case from its trace position. In Dutch, however, the topic must receive its case from the matrix verb, in the form of a preposition, a literal translation from German, as in (a) being ungrammatical[4] (in the gloss PTC stands for "particle"):

(35) a. *Deze kikker$_i$ vraagt de boer zich af ...
this frog asks the farmer himself PTC ...
... of de koning die$_i$ mag
... whether the king TP likes
"This frog, the farmer wonders whether the king likes him."

b. Van deze kikker$_i$ vraagt de boer zich af ...
of this frog asks the farmer himself PTC ...
... of de koning die$_i$ mag
... whether the king TP likes
"This frog, the farmer wonders whether the king likes him."

I will refer to the *van*-phrase as the about-argument or the about-PP. Under a traditional movement analysis the preposition *van* would have to be inserted. It is possible, but not enlightening. The analysis proposed here is that topic features move from the topic pronoun *die* to the "about" PP, *van deze kikker*.

The fact that only topics allow long distance dependencies ties in nicely with the fact that topics typically play a role not only *at* the clause level but also *above* the clause level.

Since we do not move the about PP itself but topic features of the topic pronoun, we predict that the topic pronoun must satisfy categorial selection requirements. This is correct:

(36) a. Ik vind van boeken lezen niet dat Jan dat doen moet
I find of books reading not that Jan TP:that do must
"I don't think that Jan should read books."

b. *Ik vind niet dat Jan boeken lezen doen moet
I find not that Jan books reading do should
"I don't think that Jan should read books."

(37) a. *Ik vind van boeken lezen niet ...
I find of books reading not ...
... dat mensen Jan dat moeten zien doen
... that people Jan TP:that must see do
"I don't think that people should see Jan reading a book"

b. Ik vind niet dat mensen Jan boeken moeten zien boeken lezen
I find not that people Jan books must see read
"I don't think that people should see Jan reading a book"

We have evidence both ways. The first pair shows that the topic pronoun satisfies the selection requirements in the (a) sentence, not the topic antecedent. The second pair shows that the topic pronoun fails to satisfy the selection require-

ments in the (a) sentence. The same observation were made for simple left dis-location constructions, in (15-16).

The problem with selection criteria disappears if we assume that movement targets not the topic pronoun itself but zero topic features on the topic pronoun. The pronoun satisfies or fails to satisfy the selection requirements downstairs.

Perhaps the strongest argument for our analysis is that *a bridge verb is in-dispensable for licensing island violations*. The German example in (34) and the Dutch examples following it typically involve bridge verbs. If there is no bridge verb, no island violations are possible, as shown below:

(38) a. *Jan regent het zodat  ik die nat  zie worden
         Jan rains   it   so that I  TP  wet  see  become
     b. Het regent zodat   ik Jan nat  zie worden
         It   rains  so that I  Jan wet  see  become
         "It rains so that I see Jan become wet."

*Regenen* does not select an about-constituent. However, we can also introduce such a constituent by means of sentence adverbials, and thus create a landing site for topic features:

(39) Wat Jan betreft,   het regent zodat   ik die nat  zie  worden
     As  Jan concerns it  rains   so that I  TP  wet  see  become
     "As for John, it rains so that I see him become wet."

Bridge verbs thus provide appropriate escape hatches for the topic features of the topic pronoun. The idea of escape hatches is anyhow necessary for Wh-items: there is no reason why only Wh-items should have escape hatches. Spell-out problems are avoided by moving empty topic features. Those features find their antecedent in an escape hatch, just like wh-features do.

### 4.3 Antecedents for D-pronouns

We had proposed earlier that a D-pronoun is free in the domain of its chain, and we identified the matrix sentence as that domain. The question is now: where does a D-pronoun find its antecedent? Its antecedent must be in a position where topic features are licensed. Antecedents for D-pronouns can be found in the following positions, escape hatches for topic features:

(40) a. As an "about" argument to bridge verbs
     b. In a preceding sentence
     c. As a left-dislocated constituent
     d. In an adverbial constituent of the type "with respect to X"
     e. Parenthetical
     f. *dat*-iteration construction

The latter two environments have not yet been discussed. We will first discuss parentheticals. The antecedent for a D-pronoun can also be in the other half of a parenthetical construction:

(41) Weet jij  of       die, Jan, meegaat
     know  you  whether TP,  Jan,  comes along?
     "Do you know whether Jan comes along?"

The antecedent for a D-pronoun can also be the constituent sandwiched between two complementisers in a *dat*-iteration construction (see Hoekstra 1992). The relevant construction has received little attention in the literature, because it only occurs in spoken language. If asked, linguists claim it is a processing quirk, like repititions, stuttering and so on. If that were correct then we would expect the same phenomenon to show up in other languages, like English. But in English it is systematically absent. The fact that languages differ with respect to this phenomenon suggests it is grammatical in nature. The relevant construction is given below:

(42) a. Ik denk dat  Jan dat  ik die moet vragen
        I  think that Jan that I  TP  must  ask
        "I think that John, I should ask."
     b. Ik denk dat  morgen  dat hij dan    al      komt
        I  think that tomorrow that he  TP:then already comes
        "I think that tomorrow, he'll come already."
     c. Ik denk dat  in Rome dat hij daar    graag wil   wonen
        I  think that in Rome that he  TP:there indeed wants live
        "I think that in Rome, he would like to live."

(43) a. Ik verbaas me     erover  dat Jan dat  ik die moet vragen
        I  amaze  myself about it that Jan that I  TP  must  ask
        "I'm amazed at the fact that John, I should ask."
     b. Ik verbaas me     erover  dat morgen   ...
        I  amaze  myself about it that tomorrow ...
        ... dat hij dan    al      komt
        ... that he  TP:then already comes
        "I'm amazed at the fact that tomorrow, he'll come already."
     c. Ik verbaas me     erover  dat in Rome ...
        I  amaze  myself about it that in Rome ...
        ... dat hij daar graag wil    wonen
        ... that he  there indeed wants live
        "I'm amazed at the fact that in Rome, he would like to live."

This construction is sometimes referred to as CP-recursion. This is undesirable for several reasons. CP is a bona fide syntactic category. If something like CP-recursion on the same lexical head (*dat*) exists then we would also expect AgP-recursion, VP-recursion, PP-recursion, all with the same lexical item as head. However, there is no AgP-recursion, VP-recursion or PP-recursion on the same head:

(44) a. *Ik moet moeten voetballen
      I must must play-soccer
    b. *Ik hoop op een wonder daar op
      I hope for a miracle it for

Hence we claim that the sentences above do not involve CP-recursion, but iteration.

The set of positions in which antecedents for topic pronouns are found all have in common that they are on the verge of sentential structure. None of these positions is a regular theta-position, with the possible exception of the argument to a bridge verb. Note though that the bridge verb does not assign a normal theta role to the argument in the "about" phrase. The "about" phrase is not a participant in the action described by the bridge verb like the subject is, nor is the "about" phrase brought about by the action described by the bridge verb. In German, the about-phrase shares morphological case with the topic pronoun. In Dutch, the default preposition *van* shows up.

We could suggest that the set of environments given above all share the common feature that the positions are all escape hatch positions through which the larger structure in which the sentence occurs is accessed. We could then maintain the claim that topic pronouns are free in the matrix sentence. It would lead us too far afield to try to work out such a proposal.

## 5. Concluding Remarks

In this explorative article the distribution of the topic pronouns *die/dat* was discussed. It was shown that these pronouns have to be free in the domain of the matrix sentence. It was also shown that they can find their antecedent in a set of positions which we characterised as escape-hatches for topic features. This allowed us to relate a topic pronoun and its antecedent by means of the movement of zero topic features. We will leave it to future research to further develop the idea of viewing movement of topic features as a way of formally characterising the relation of coreference between a topic pronoun and its antecedent.

## Address of the Author

Eric Hoekstra
Frisian Academy
Royal Academy
P.O. Box 54
NL-8900 AB Ljouwert/Leeuwarden (The Netherlands)

## Notes

1    I would like to thank Marjo van Koppen, Jarich Hoekstra and an anonymous reviewer for comments and/or discussion. According to the latter, *die* in (1b) initiates a new full clause,

74

while the coreferential *De minister* is left-extraposed. I contest this insight on the basis of my native competence of Dutch: (1b) is one single clause just as well as (1a).

2    Prescriptive grammar states that in referring to a *het* noun, *dat* must be used, and in referring to a *de* noun, *die*. As shown in Romijn (1996), this is not a description of the facts.

3    There is also a possessive topic pronoun *diens* (see Postma 1984). It is typical of written language. In spoken language an analytic form is used, namely *die z'n*, which can be glossed as "TP his".

4    The preposition could remain absent in Middle Dutch (Van den Berg 1992), as shown below in (i):

(i) der Walewein     claechde     sijn swert dat hijt   daer niet en hevet
    the lord Walewein complained his   sword that he-it there not  has
    "The lord Walewein complained that he didn't have his sword there."

Thus Middle Dutch resembles German more than Modern Dutch.

# References

Berg, E. van den. 1992. "Zinsvermenging in het Middelnederlands." Taalkundig Bulletin 22, 101-108.

Chomsky, N. 1982. Some concepts and consequences of the theory of government and binding. Cambridge, Mass.: MIT Press.

Chomsky, N. 1995. The Minimalist Program. Cambridge, Mass.: MIT Press.

Grohmann, K. 1997. "On Left Dislocation." In W. Abraham & K. Grohmann (eds.), Groninger Arbeiten zur germanistischen Linguistik (GAGL) 40, 1-33.

Heim, I. 1982. The Semantics of definite and indefinite Noun Phrases. New York: Garland.

Hoeksema, J. & D. J. Napoli. 1993. "Paratactic and subordinative *So*." Journal of Linguistics 29, 291-314.

Hoekstra, E. 1992. On the Parametrization of Functional Projections in CP. In: A. J. Schafer (ed.), 23th Proceedings of the North-East Linguistic Society 1992. Amherst: University of Massachusetts, 191-204.

Hoekstra, E. & J.-W. Zwart. 1994. "De Structuur van de CP. Functionele Projecties voor Topics en Vraagwoorden in het Nederlands." Spektator 23, 191-212.

Hoekstra, J. 1995. "Preposition stranding and resumptivity in West-Germanic." In: H. Haider, S. Olsen & S. Vikner (eds.), Studies in Comparative Germanic Syntax (=Studies in Natural Language and Linguistic Theory, vol. 31). Kluwer, Dordrecht, 95-118.

Postma, G. 1984. "The Dutch pronoun *diens*; distribution and reference properties." In: H. Bennis & W.U.S. van Lessen Kloeke (eds.), Linguistics in the Netherlands 1984. Dordrecht: Foris.

Reinhart, T. 1981. "Pragmatics and Linguistics: an Analysis of Sentence Topics." Philosophica 27, 53-94.

Reinhart, T. 1983. Anaphora and Semantic Interpretation. London: Croom Helm.

Reinhart, T. & E. Reuland. 1991. "Anaphors and Logophors: an Argument Structure Perspective." In J. Koster & E. Reuland (eds.), Long Distance Anaphora. Cambridge: Cambridge University Press, 183-321.

Romijn, K. 1996. Hoe doen we het? Verwijzen naar linguïstische en cognitieve representaties met het voornaamwoord het. Amsterdam: Publicaties van het Meertensinstituut.

Vries, W. de. 1910-1912. Dysmelie. Opmerkingen over syntaxis. Verhandeling behorende bij het programma van het gymnasium der gemeente Groningen voor het jaar 1910-1911. Idem voor het jaar (vervolg) 1911-1912. Groningen: B. Jacobs.