On the status of agreement and relative clauses in West-Flemish

HANS BENNIS

Free University of Amsterdam

LILIANE HAEGEMAN

University of Ghent

Introduction

It is well-known that, from a typological point of view, relative clauses are extremely diversified across languages. English, Dutch, and French, for example, differ in several ways with respect to relativization, and the picture becomes even more varied if we include a wider range of languages.

In this article we shall first look at the relative clause construction in one particular language: West-Flemish. Relativization in West-Flemish will be shown to differ in rather crucial aspects from that in surrounding languages like English, Dutch, and French. Of course, it would be possible to devise mechanisms capable of providing an adequate description of the facts of West-Flemish relativization within a particular linguistic framework. However, such an approach would only achieve descriptive adequacy, and would fail to give us any further insight into the structure of human language, i.e. it will lack explanatory adequacy. In order to obtain a level of explanatory adequacy, we must determine how a specific set of facts in a language is related to other properties of that language, and, more generally, we should look for those correlations which may lead to interesting generalizations and finally to an explanation of the phenomena observed in terms of the fairly abstract system of Universal Grammar.

Universal Grammar (UG) is conceived of as an abstract representation of the concept of "possible grammar of a natural language". It can be viewed as a system of general rules and principles, some of which are parametrized to allow for diversification across languages. The core grammar of a specific language, say English, is determined by fixing the parameters left open in UG. In order to gain optimal explanatory adequacy, it is desirable to restrict the number of parameters. (cf. Chomsky 1981, 1982).

In this article the observations concerning West-Flemish relative clau-

ses which at first sight might appear rather problematic for the analysis will be shown to be a consequence of other more general properties of the language. The relevant properties of West-Flemish will be seen to be related to independently motivated principles of UG. More specifically, we shall claim that the fact that West-Flemish is a pro-drop language – i.e. a language which allows the subject pronoun to be dropped under specific conditions – will account for the patterns observed in the case of relative clauses. The pro-drop parameter can be assumed to be part of UG: it determines the typological diversification across languages of a set of correlating properties, as has been described in the extensive literature on this topic within the generative framework. (cf. Chomsky 1981, Ch. IV; Chomsky 1982, Rizzi 1982, Ch. IV; Safir 1982 b, Ch. VI)

This article is organised as follows. In the first part we shall identify the problem posed for the description by relative clauses in West-Flemish. Our analysis will lead us to the conclusion that by applying the devices available for the analysis of relativization in other languages we would end up with a set of unattractive ad-hoc stipulations.

In the second section we shall discuss the phenomenon of complementizer agreement. This will lead to the conclusion that West-Flemish has the relevant properties of a pro-drop language. For reasons of space, this section will be rather short and incomplete. For an exhaustive discussion of the facts and their consequences for the theory, we refer to a forthcoming publication (Bennis & Haegeman, forthcoming).

In the last section of this article, we shall discuss the consequences of our analysis of complementizer agreement for the discussion of relativization in West-Flemish. This section will show how our analysis of the agreement phenomena in West-Flemish also allows a rather simple and elegant explanation for what at first appeared to be rather problematic facts with respect to relativization. The ad-hoc stipulations which at first seemed necessary will be shown to be dispensable.

Relative Clauses I

The language considered here is a Flemish dialect spoken in the rural area of West-Flanders situated between Knokke-Heist and Bruges. In this dialect a clear and unexpected subject-object asymmetry emerges in relative clauses. The general pattern is indicated in (1) & (2):

- (1) (a) ... den vent da Pol t getrokken heet
 (b) *... den vent die Pol t getrokken heet
 ... the man that/whom Pol t has made a picture of
- (2) (a) ... den vent da t gekomen is
 (b) ... den vent die t gekomen is
 ... the man that/who t has come

In (1) the object of the relative clause is relativized. In this case only the lexical complementizer *da* (1a) and not the relative pronoun *die* (1b) can

follow the head noun *vent*. If the subject is relativized, as in (2), both the lexical complementizer (2a) and the relative pronoun (2b) lead to a grammatical result. The same subject-object asymmetry can be observed with long relativization as in (3) and (4):

- (3) (a) ...den vent da Pol peinst *da* Marie t getrokken heet
(b) *...den vent da Pol peinst *die* Marie t getrokken heet
...the man that Pol thinks that/whom Marie has made a picture of
- (4) (a) ...den vent da Pol peinst *da* t gekomen is
(b) ...den vent da Pol peinst *die* t gekomen is
...the man that Pol thinks that/who has come

Surprisingly, the subject-object asymmetry in these sentences shows up in the most embedded sentence or, to be more precise, it shows up in the sentence from which the relative pronoun is extracted.

The observations with respect to object relativization lead to a clear generalization: the relative pronoun never shows up. From this we might conclude that the relative pronoun is obligatorily deleted under identity with the adjacent head of the relative construction, after application of *wh*-movement of the relative pronoun to the COMP-position. So we might think of the following derivational stages for a sentence like (3a):

- (5) (a) ...den vent da Pol peinst da Marie *die*
getrokken heet (D-str)
(b) ...den vent da Pol peinst *die_i* da Marie *t_i*
getrokken heet (*wh*-mvt)
(c) ...den vent *die_i* da Pol peinst *t_i* da Marie *t_i*
getrokken heet (*wh*-mvt)

(5c) is the stage prior to the application of the deletion of *die* under identity with the adjacent head (*den vent*).

There is no reason to assume that the derivation of subject relativization will be different. So for (4) we will have the corresponding derivational stages given in (6):

- (6) (a) ...den vent da Pol peinst da *die*
gekommen is (D-str)
(b) ...den vent da Pol peinst *die_i* da *t_i*
gekommen is (*wh*-mvt)
(c) ...den vent *die_i* da Pol peinst *t_i* da *t_i*
gekommen is (*wh*-mvt)

However, applying deletion under identity to (6c) will only give us the pattern of (4a). The question then arises how to account for the occurrence of a sentence like (4b), with *die* rather than *da* in the lowest COMP-position.

In trying to answer this question it might seem useful to observe that (4b) exhibits a pattern which resembles that of French relative clauses. As is well-known French relatives also show a subject-object asymmetry which is remarkably similar to the West-Flemish pattern. The relevant data are given in (7) and (8). With object relativization only the lexical complementizer *que* shows up, while in the case of subject relativization *qui*, the relative pronoun, must appear:

- (7) (a) ...l'homme *que* j'aime t
 (b) *...l'homme *qui* j'aime t
 ...the man that/whom I love t
- (8) (a) *...l'homme *que* t viendra
 (b) ...l'homme *qui* t viendra
 ...the man that/who t will come

Note, however, that the distribution of lexical complementizer and relative pronoun in French is not identical to that in West-Flemish. As (8a) shows the relative pronoun *qui* is obligatory in the case of French subject relatives, while West-Flemish subject relatives have either *die* or *da* (cf. 2).

The same subject-object asymmetry appears with long relativization in French, as can be seen in (9) and (10):

- (9) (a) ...l'homme *que* tu crois *que* j'aime t
 (b) *...l'homme *que* tu crois *qui* j'aime t
 ...the man that you think that/whom I love t
- (10) (a) *...l'homme *que* tu crois *que* t viendra
 (b) ...l'homme *que* tu crois *qui* t viendra
 ...the man that you think that/who t will come

Sentence (10b) is parallel to the West-Flemish example (4b). But the ungrammatical sentence (10a) is parallel to the grammatical West-Flemish sentence (4a). Hence, if we want to explain the patterning of relative clauses in West-Flemish along the lines of the explanation given for the French relative construction, the question will arise of how to explain the occurrence of a sentence like (4a). More importantly, it will turn out that an analysis of the West-Flemish data based on the account given for French relatives will lead to a number of unattractive ad-hoc stipulations.

With respect to French relatives it is argued that the lexical complementizer *que* is changed into a kind of relative pronoun, or, to be more precise, an indexed complementizer, *qui*, whenever it is followed by a nominative trace, i.e. in case of subject extraction. Pesetsky (1982) states the rule as follows:

- (11) $[\text{COMP} \left\{ \begin{smallmatrix} \text{WH}_i \\ t_i \end{smallmatrix} \right\} \text{que}] \rightarrow [\text{COMP } \text{qui}_i] / \text{---} [\text{s} \times [t_{i+nom}]_y]$

Although this rule is language-specific, and therefore not very interesting from a theoretical point of view, it is motivated by some independent and more general principles. The main motivation of the rule relates to the structural asymmetry between subject and object.

Within the Government-Binding framework (Chomsky 1981) there exists a requirement to the effect that traces must be locally identified. This requirement is stated in the form of the Empty Category Principle (ECP) which demands that traces be properly governed. With respect to the fulfilment of the ECP subject traces differ from object traces, in that the latter are properly governed by the governing verb, while the former will only be properly governed if they can be locally identified by the moved constituent or its trace in the COMP-position.

In addition, French has the requirement that the COMP-position contain only one constituent, the so-called doubly filled COMP filter (cf. Chomsky and Lasnik 1977).

These two requirements, ECP and doubly filled COMP filter, motivate the application of rule (11). When the subject is relativized, its trace must be properly governed by coindexing (ECP). On the other hand, the doubly filled COMP filter will force the deletion under identity of the *wh*-word or its trace in COMP, thus in fact preventing that the subject position is properly governed, and causing a violation of the ECP. In order to avoid a violation of the ECP, French must make use of the language specific rule (11) which will guarantee that both the doubly filled COMP filter and the ECP can be met, since the single filler of the COMP position after the application of (11) is the coindexed complementizer *qui*_i, which will properly govern the nominative trace (*t*_i). In cases of subject relativization rule (11) is therefore obligatory.

In the case of object relativization, on the other hand, a rule like (11) is not needed, since the subject trace will be properly governed by the verb.

Though language specific, rule (11) in French can thus be seen to be motivated by independent and more general principles of the grammar.

For the description of the West-Flemish data under discussion it might be feasible to devise a rule along the lines of (11) to account for the occurrence of the relative pronoun *die* with subject relatives.¹ However, this rule must be optional in order to allow for both possibilities of subject relativization (cf. (4a) and (4b)). From the optionality of the rule it follows that the ECP cannot provide an independent motivation for such a rule, as it does in French: as (4a) shows, a rule like (11) is not necessary in order to meet the ECP. Any West-Flemish equivalent of rule (11) will therefore be a purely descriptive device, which lacks any independent motivation.

An alternative approach might be to argue that with respect to subject relativization West-Flemish exhibits the pattern seen in English relative clauses, which indeed show an alternation between a complementizer and a relative pronoun, as can be seen in (12).

- (12) (a) ...the man that/who/e I saw t
 (b) ...the man that/who/*e t saw me
- (13) (a) ...the man that/who/e I think that/*who/e I saw t
 (b) ...the man that/who/e I think *that/*who/e t saw me

As has been shown by Pesetsky (1982) this distribution can be explained by the interaction of (i) the independently motivated ECP which explains the ungrammaticality of *that*-trace sequences,² (ii) the doubly filled COMP restriction, which is also found in French, together with (iii) an optional deletion of constituents in COMP. In order to describe the distributional pattern in (12)–(13) only one specific rule is needed which is thus independently motivated, as was rule (11) in French.

An analysis of West-Flemish on the basis of the analysis provided for English by Pesetsky is equally implausible, though. First, the alternation between complementizer (*that*) and relative pronoun (*who*) in English, is not dependent on a subject-object asymmetry (12), while the corresponding alternation in West-Flemish clearly shows such an asymmetry ((1) and (2)).

Additional complications for any explanations based on the English data are caused by the observation that West-Flemish in general seems to allow complementizer trace sequences,³ it allows the COMP-position to be doubly filled, and it does not allow the complementizer to be deleted, with the sole exception of the subject relativization cases. These three observations are illustrated in (14)–(16).

- (14) Wien peinst Pol da t Valère gezien heet?
 Who thinks Pol that t Valère seen has
- (15) Ik weten niet wien da Pol t gezien heet
 I know not whom that Pol t seen has
- (16) *Ik weten niet wien Pol t gezien heet
 I know not whom Pol t seen has

In (14) the subject of the embedded clause is questioned. The pattern resulting from the application of *wh*-movement is the standard case of a complementizer trace sequence which is ungrammatical in English. In West-Flemish, as indeed in other variants of Dutch, these sentences are grammatical (cf. Maling and Zaenen (1978); Bennis (1980); Koopman (1983)).

In (15) the COMP position contains both the moved *wh*-constituent *wien* and the lexical complementizer *da*. This is the normal pattern in West-Flemish; in fact, as (16) shows, deletion of the lexical complementizer leads to ungrammaticality. The complementizer is only deleted in the case of subject relativization in West-Flemish, as is seen in (2b) and (4b).

Summarizing this section, we have demonstrated that West-Flemish relative clauses show a subject-object asymmetry, which cannot be ex-

plained adequately by means of the types of analyses given for French or English.⁴

Though it is perhaps possible to formulate some descriptively adequate device to account for the phenomena, such an ad-hoc device cannot be motivated since West-Flemish lacks the restriction on doubly filled COMP and the rule of optional deletion in COMP. Any such ad-hoc account of the facts will thus be a mere stipulation without any explanatory value. Since this is a very unsatisfactory conclusion, we shall try to present a radically different solution based on the analysis of a seemingly unrelated set of facts in West-Flemish.

West-Flemish as a pro-drop language

In this section we shall argue that a set of facts in the West-Flemish dialect at issue show a remarkable resemblance with patterns relating both to pro-drop phenomena and to the process of cliticization and clitic doubling. These patterns are familiar from the literature on Romance languages (Jaeggli 1982, Rizzi 1982). The discussion of the West-Flemish data will enable us to treat both the pro-drop phenomena and the phenomena related to cliticization and clitic-doubling in terms of one generalized account of the possible occurrence of empty pronominal noun phrases. After having provided a global analysis for the facts, which at first sight look rather problematic, we shall return to the subject-object asymmetry as discussed in the previous section.

In West-Flemish the lexical complementizer which introduces finite sub-ordinate clauses shows obligatory agreement in number with the subject of the clause it introduces. In (17), for instance, the complementizer *da* is singular as it introduces a sentence with a singular subject; in (18) the complementizer *dan* is plural, since the subject *Pol en Valère* is plural.

- (17) ...*da* Pol zat is
 ...that Pol drunk is
 (18) ...*dan* Pol en Valère zat zijn
 ...that Pol and Valère drunk are

The difference in form of the complementizers in (17) and (18) suggests that these complementizers should not only be specified for the feature $\pm$ Tense (depending on whether or not they introduce finite clauses), but that the complementizer of finite clauses (*da(n)*) also has to be specified for the feature $\pm$ plural, where the selection of + or – depends on the number of the subject NP of the clause introduced by *da(n)*. The grammatical representation of the complementizer will be something like (19):

- (19) *da*_[+Tense, –plural] *dan*_[+Tense, +plural]

In order to deal with these facts within the framework of the Government-

- (23) INFL $\rightarrow \pm$ Tense – AGR
 AGR = [+N, –V], [α number, β person, γ gender]

With respect to the position of INFL in V_2 -languages, two analyses are available. The first option is that already discussed for English and Italian and given in (22) (cf. Safir (1982a). and others): INFL is part of S.

Alternatively it has been argued that INFL is part of COMP; this approach is supported by the fact that COMP must be specified for $\pm$ Tense anyway (cf. 20). This latter analysis suggests that we modify (20) to (24):⁷

- (24) COMP $\rightarrow$ XP – INFL

On the basis of (24) the obligatory agreement in West-Flemish of the lexical complementizer with the subject of the clause it introduces can be accounted for in a straightforward way. The position INFL contains AGR with its pronominal features, which include number; hence we shall postulate that the lexical complementizer actually results from spelling out the features of INFL. West-Flemish complementizers in finite clauses have to be specified for the feature number, in addition to the feature [+Tense]. If AGR (in INFL, in COMP) contains the feature [–plural], as in (17), the complementizer will be spelt out as *da*.

Pursuing this approach somewhat further, we might then ask what would happen if all pronominal features of AGR were to be specified. In fact, such a full specification should actually result in a complementizer which agrees in number, person and possibly gender with the subject NP. AGR would thus presumably end up as some kind of pronoun: it is nominal [+N, –V] and now also fully specified for all pronominal features. This is exactly what can be observed in West-Flemish, as is indicated in (25) which gives us a subordinate clause with a pronominal subject:

- (25) ...dase zie komt⁸
 ...that[–plu, 3pers, +fem] she comes

The full paradigm for pronominal subjects is given in (26):

(26)	...dank (ik) kommen	sing, 1 pers
	...daj (gie) komt	sing, 2 pers
	...datje (jij) komt	sing, 3 pers, masc
	...dase (zie) komt	sing, 3 pers, fem
	...dat (et) komt	sing, 3 pers, neut
	...dame (wunder) kommen	plu, 1 pers
	...daj (gunder) komt	plu, 2 pers
	...danze (zunder) kommen	plu, 3 pers

The round brackets around the pronoun in subject position indicate that it is optional. The actual appearance of subject pronouns in these sentences is somewhat marked in that they only occur if the subject receives stress. This phenomenon is in line with the Avoid Pronoun Principle which predicts that the pronoun will be dropped whenever possible (cf. Chomsky, 1981). Since an inflected complementizer cannot be stressed adequately, stress on the subject will be a sufficient reason for retaining the subject pronoun.

It might perhaps be possible to analyse the sentences of (26) without the overt subject pronoun as instances of cliticization of the subject pronoun to COMP. This phenomenon is to be found in most dialects of Dutch.⁹

(27) ... da [se]_i t_i komt

However, if we were to analyse cliticization as a movement rule, i.e. movement of the subject pronoun to the right of COMP, as in (27), it would be difficult to relate this to the obligatory agreement between complementizer and non-pronominal subject in cases like (17)–(18), and, secondly, an additional reduplication rule would be necessary for sentences like (26) with the overt subject pronoun.

In several respects, of course, the *se* part of *dase* is clitic-like. It is nominal, it contains pronominal features for number, person and gender, and it cannot be stressed. If we analyse this clitic-like element as being base generated in COMP, i.e. as the result of spelling out the features of AGR in INFL, our analysis will in most aspects be similar to the analysis of object clitics in Romance languages, as proposed by Jaeggli (1982) and others. In this approach the object clitic is base generated in its surface position and related to a base generated empty NP in argument position. It is known that languages like Spanish allow this argument position also to be filled by a coreferential pronoun, if the object pronoun is stressed, as indicated in (28).

- (28) (a) Lo_i vimos [e]_i
 (b) Lo_i vimos a el_i
 (We) see him

It will be clear from a comparison of (26) in West-Flemish with the Spanish data in (28) that there is a striking similarity between the COMP-agreement phenomena and the facts of object clitics and clitic-doubling.

But since it is a subject pronoun that can be dropped in West-Flemish, the facts also are relevant to the pro-drop parameter. Several languages allow the subject pronoun to be dropped, a phenomenon which depends on the richness of the inflectional paradigm of the verb. Within the Government-Binding framework it is assumed that it is specific properties of INFL that determine the variation among languages with respect to the pro-drop phenomena. In Italian, which has a relatively rich inflectional

paradigm, the node INFL is sufficiently specified, allowing the subject pronoun to be dropped, while the subject pronoun cannot be dropped in English. Several proposals have been made to account for these pro-drop phenomena within the theory. For the moment we will accept Chomsky's most recent proposal (Chomsky, 1982). In contrast with previous analyses, Chomsky argues that the empty subject in finite sentences in languages like Italian differs in status from the empty subject of infinitivals. Infinitivals are taken to have as their empty subjects PRO, the pronominal anaphor which only appears in ungoverned positions. In previous accounts the empty subject of finite sentences was also assumed to be PRO, i.e. ungoverned, but in Chomsky (1982), Chomsky argues that empty subjects in finite clauses should be seen as empty pronominals occurring in a position governed and hence assigned nominative case by AGR. This empty pronominal is indicated by pro:

- (29) pro = [+ pronominal, -anaphor]
 governed by AGR
 assigned nominative case by AGR

Hence the pro-drop parameter can be formulated in terms of some variable property of AGR which determines the possible occurrence of pro as the subject of finite clauses. More precisely, the content of pro has to be determined by agreement with its governor, AGR. The absence of some crucial property of AGR in the non-pro-drop languages blocks the proper determination of the content of pro. Chomsky suggests that the pro-drop parameter might be defined on the presence or absence of Case on AGR in D-structure. Another possibility, which is developed on the basis of the West-Flemish evidence, is suggested here.

With respect to West-Flemish complementizers we proposed that clitic-like element *-se* which is part of the complementizer *dase* in (25) might be analysed as the spelling out of all the pronominal features of AGR (in INFL, in COMP). In the presence of this element *-se* the subject position itself may be left empty. It is thus obvious that this empty subject position has the properties of a subject in a pro-drop language. Assuming our discussion of the pro-drop parameter above, we can say that the subject in West-Flemish can be pro because the content of pro is properly determined by agreement with AGR, the latter being here specified for all relevant pronominal features. In Italian the inflectional paradigm which guarantees full determination of pro, is always that of the verbal inflection. In West-Flemish the relevant inflectional paradigm which allows pro as a subject shows up as complementizer inflection in subordinate clauses. This difference relates to the fact that West-Flemish is a Verb Second language.¹⁰

At this point we must point out that not any agreement relation between AGR and the subject position will be sufficient to determine the content of pro in subject position. We have seen that overt agreement

between complementizer and subject is obligatory in West-Flemish, but if this agreement only concerns number then the subject position cannot be *pro*:

- (30) *...da *pro* komt
 ...that_[-plu] *pro* comes
 *...dan *pro* kommen
 ...that_[+plu] *pro* come

This leads us to the conclusion that *pro* will only be allowed to occur if its governor is specified for all pronominal features. The governor of *pro* will be AGR for subject *pro* and a clitic for object *pro*. In West-Flemish AGR is either specified for the pronominal feature number only, or for all the pronominal features, i.e. person, number, gender as indicated in (31):

- (31) a
INFL	
+T	$\left[\begin{array}{c} -\text{plu} \\ 0 \text{ pers} \\ 0 \text{gend} \end{array} \right] \rightarrow da$
	AGR
- b
INFL	
+T	$\left[\begin{array}{c} -\text{plu} \\ 3 \text{ pers} \\ +\text{fem} \end{array} \right] \rightarrow dase$
	AGR

This analysis of empty pronominals leads us to a global formulation of the pro-drop parameter which will now be taken as a general principle that determines the distribution of empty pronominal argument positions, i.e. both subject and non-subject positions can be subsumed under it.

- (32) pro-drop parameter
 a language allows *pro* in argumentposition if the pronominal properties of *pro* are properly determined by the governor of *pro*

We shall now propose a more formal account of this basic idea.

First we have to establish the conditions under which the proper determination of *pro* takes place. We propose that *pro* is properly determined if all its pronominal features are specified by agreement with AGR or Clitic. Consequently, we shall assume that in Italian, a pro-drop language, all pronominal features of AGR are specified, thus creating the possibility of the occurrence of *pro* in the subject position. In English, on the other hand, AGR is not inherently specified at all, and the introduction of *pro* in subject position would leave the content of *pro* completely undetermined.

A familiar way to indicate the agreement relation between AGR and the subject position is to introduce a superscripting convention. We shall use the superscripting convention for both the relation between AGR and *pro* and that between clitic and *pro*.¹¹ The convention of coindexation by superscripting is defined as in (33):

- (33) Superscripting Convention
 α is coindexed with β by superscripts
 iff- α governs β
 $-\alpha = [+N, -V]$
 $-\alpha$ contains at least one specified pronominal feature
 $-\beta$ contains one or more pronominal features

From this formulation it follows that in a language which has clitics or a (partly) specified AGR, there may exist a superscript relation between a clitic or AGR and a corresponding argument position.¹²

If we now apply this definition to the complementizer agreement phenomena of West-Flemish, we see that AGR in INFL in COMP will always be cosuperscripted with the subject position, since it is obligatorily specified for the feature number, a pronominal feature. This can be seen in (34) and (35):

- (34) (a) ...daⁱ Polⁱ komt (that_[-plu] Pol comes)
 (b) *...daⁱ zieⁱ komt (that_[-plu] she comes)
 (c) *...daⁱ proⁱ komt (that_[-plu] pro comes)
 (35) (a) *...daseⁱ Marieⁱ komt (that_[-plu, 3 pers, fem] Marie comes)
 (b) ...daseⁱ zieⁱ komt (that_[-plu, 3 pers, fem] she comes)
 (c) ...daseⁱ proⁱ komt (that_[-plu, 3 pers, fem] pro comes)

We have shown before that if the complementizer is specified for the feature number only, as in (34), non-pronominal NP's appear in subject position (34a) and pro is excluded (34c). Surprisingly perhaps, it is also impossible for a pronoun to appear in subject position (34b). In other words, all pronominals (pro and pronoun) are excluded in this configuration.

Pronominals appear in subject position only if AGR is fully specified, as in (35). Conversely the insertion of a non-pronominal NP in subject position in (35) leads to ungrammaticality (35a).

In both paradigms (34) and (35) pro and pronoun thus pattern together, and contrast with non-pronominal NP's. Apart from the consideration that this provides us with an argument for treating the empty subject position in (35c) as an empty pronominal (pro), it also forces us to relate these phenomena to a principle which distinguishes pronominals from non-pronominals. Within the Government-Binding framework there are at least two such principles available. First, there is the Binding Theory: pronominals are subject to principle B, non-pronominals are subject to principle C (cf. Chomsky 1981). A second principle, concerned solely with pronouns is the Avoid Pronoun Principle (cf. Chomsky 1981). With respect to (34) and (35) the Avoid Pronoun Principle explains the relative markedness of the grammatical (35b), but it does not explain the ungrammaticality of (34b) and (35a). Neither is it quite clear at this stage how the grammaticality judgements of (34) and (35) could follow directly

from the binding theory or from an extension of it. This leaves us with the rather unfortunate conclusion that we shall have to develop yet another principle which distinguishes pronominals from non-pronominals. Although eventually this principle might be shown to follow from a more general account of the differences between pronominals and non-pronominals, we shall here develop a rather specific principle and show its consequences with respect to a wide range of data from the different languages.

The principle is formulated as a condition on superscript relations, in contrast to the Binding Theory which might be considered as a condition on subscript-relations. Before we formulate the principle we have to introduce an assumption which distinguishes between pronominals and non-pronominals. We consider a pronominal as a set of specified pronominal features with or without a phonological matrix. We further assume that non-pronominal NP's are only grammatically specified for the pronominal feature number, and not for person and gender. This assumption is based on the observation that the feature person is redundant: non-pronominals are third person; and that the feature gender is best analysed as a lexically or pragmatically determined feature of the head of the non-pronominal NP. Only the feature number has a clear grammatical function.

In order to account for the distributional pattern indicated in (34) and (35), we propose a principle that requires that cosuperscripted elements be maximally identical. It is clear that some such principle will be needed, since the grammar will have to indicate somewhere that the singular complementizer *da* in (34a) cooccurs with a singular subject, requiring thus that both features number, of complementizer and subject NP, have the same value. The principle will be formulated as a Maximal Identity Principle:

- (36) Maximal Identity Principle
 α is maximally identical with β
 iff for each pronominal feature γF_i in α , β contains a feature δF_i
 such that $\gamma = \delta$ ($\gamma, \delta = +, -, \text{ or } 0$)

Assuming such a principle, we can explain the different patterning of pronominals and non-pronominals in (34) and (35):

- (37)
INFL	NP	NP
$+T$ $\left[\begin{array}{l} -\text{plu} \\ 0 \text{ pers} \\ 0 \text{ gender} \end{array} \right]_{\text{AGR}}$	$\left[\begin{array}{l} -\text{plu} \\ 3 \text{ pers} \\ +\text{fem} \end{array} \right]$	$\left[\begin{array}{l} -\text{plu} \\ 0 \text{ pers} \\ 0 \text{ gender} \end{array} \right]$
<i>da</i>	* <i>zie</i> /* <i>pro</i>	Pol

In (37), which corresponds to (34), AGR is only specified for the feature number and not for person and gender. Therefore it is maximally identical to a non-pronominal NP. The ungrammaticality of pronoun or pro follows from the principle since these pronominals are not identical to AGR. Note that we have subsumed the pro-drop parameter as an instance of this principle. From (36) it follows that pro may appear only if AGR is fully specified, which allows the pronominal properties of pro to be properly determined.

The ungrammaticality of (35a) follows from (36) too: the Maximal Identity Principle is not satisfied. Hence (36) will account for both (34) and (35). This principle (36) should be seen as a condition on agreement.

In Bennis and Haegeman (forthcoming) we shall argue that the principle is able to provide us with an explanation of the difference between pronominals and non-pronominal NP's with respect to clitic doubling in Standard Spanish, a phenomenon not accounted for in Jaeggli's analysis (1982). The principle will also be shown to account for the fact that Welsh shows only pronominal agreement (cf. Harlow 1983), and for the inflection phenomena in Modern Irish (cf. McCloskey and Hale 1983).

This theory of Agreement, with the superscripting convention (33) and the Maximal Identity Principle (36), allows us to account for a wide range of agreement phenomena, such as subject-verb agreement, NP-internal agreement (Int 12), pro-drop and clitic phenomena, and the contrast between pronominal and non-pronominal agreement. For reasons of space we restrict our attention in this article to the discussion of the relevant West-Flemish data.

We have already discussed the general pattern of complementizer agreement in subordinate clauses. We have argued that West-Flemish is a verb-second language with INFL in COMP. In main clauses we expect more or less the same pattern as in subordinate clauses with respect to the agreement phenomena. The main difference with subordinate clauses is that the lexical complementizer is absent and that agreement will show up on the finite verb which has been moved into the +Tense position. If the XP position in COMP (cf. 24) is occupied by a constituent which is not the subject of the main clause, the pattern that emerges mirrors that for subordinate clauses (cf. (34) & (35)):

- (39) (a) Dat geld heet Marie hem gegeven
 (b) *Dat geld heet zie hem gegeven

- (c) *Dat geld heet pro hem gegeven
That money has_[-plu] Marie/*she/*pro him given
- (40) (a) *Dat geld heese Marie hem gegeven
(b) Dat geld heese zie hem gegeven
(c) Dat geld heese pro hem gegeven
That money has_[-plu, 3 pers, +fem] *Marie/she/pro him given

These data are in all relevant respects similar to (34) and (35).

If the subject of the main clause is moved into the XP-position in COMP, the situation changes quite drastically:

- (41) (a) *Marie_i goase t_i kommen
(b) *Zie_i goase t_i kommen
(c) *pro_i goase t_i kommen
*Marie/*She/*pro goes_[-plu, 3 pers, +fem] come
- (42) (a) Marie_i goat t_i kommen
(b) Zie_i goat t_i kommen
(c) *pro_i goat t_i kommen
Marie/She/*pro goes_[-plu] come

Apparently the grammaticality of (42b) and the ungrammaticality of (41b & c) violate the theory of agreement developed so far. Note however that the superscript relation in (41) between AGR and the subject position is one between AGR and a trace. This trace is a variable, left by the topicalization of the subject. As is familiar from Chomsky (1981) traces of *wh*-movement type operations, or traces that are $\bar{A}$ -bound, are variables and variables behave like names. Names are non-pronominal and consequently the maximal identity principle will be violated in (41) in the same way as it is violated in (40a).¹³

In order to account for agreement in passive or subject-raising sentences we have to assume that the agreement relation between AGR and the subject position will be achieved at the level of S-structure; hence there will be no cosuperscripting relation between the topicalized subject itself and AGR in (41) and (42), but only between AGR and the trace in subject position. What this means is that *pro* cannot appear in XP, since it will be undetermined by the lack of agreement with AGR, there being no cosuperscripting relation. This explains the ungrammaticality (42c). (42a) and (42b), on the other hand, are grammatical: pronouns and non-pronominal NP's need not be properly determined by AGR (and indeed they are not), and since there is no cosuperscripting between the topicalized subject and AGR there is no maximal identity requirement, hence both pronouns and non-pronominal NP's are allowed. We thus see that *pro* and pronoun pattern together if they occur in a configuration in which they are cosuperscripted with another element for agreement; if not, lexical pronouns behave like non-pronominal NP's in this respect.

Finally, a very interesting set of facts is found in (43):

- (43) (a) *Ze-goat Marie kommen
 (b) Ze-goat zie kommen
 (c) Ze-goat pro kommen
 She (weak form) goes *Marie/she/pro come

Although (43c) may at first sight look like (42b) with a weak rather than a strong subject pronoun, we shall argue that in (43) it is not the subject that has been moved to the XP position in COMP, but rather the AGR-clitic. This approach is confirmed by the grammaticality of (43b) where the strong pronoun *zie* must be analysed as occurring in the subject position, since the clitic-like element in AGR is *-se* rather than *zie* (cf. 41). If we analyse (43) as involving a movement of the clitic-like element in AGR to XP in COMP, we shall expect pronoun and *pro* to pattern together: after movement there will still be an agreement relation between the moved AGR and the subject position, and the maximal identity principle will have to apply, ruling out nonpronominal NP's in subject position (43a).

In this section, we have thus argued that West-Flemish provides arguments for a generalized account of the *pro*-drop and clitic phenomena.¹⁴ We have made our account explicit by the introduction of a superscripting convention and a condition on superscripted constituents: the Maximal Identity Principle. These principles enabled us to account for a wide range of data. In the last section of this article we show what the account provided thus far can achieve with respect to the problematic relativization phenomena in West-Flemish as discussed in a previous section.

Relative Clauses II

On the basis of the properties of COMP and the status of Agreement in West-Flemish, as discussed in the preceding section, the relativization data introduced in the beginning of this article can be accounted for without additional ad-hoc assumptions. The facts to be explained are repeated here as (44)–(47).

- (44) (a) ...den vent da Pol t getrokken heet (= 1a)
 (b) *...den vent die Pol t getrokken heet (= 1b)
 ...the man that/whom Pol t has made a picture of
- (45) (a) ...den vent da t gekomen is (= 2a)
 (b) ...den vent die t gekomen is (= 2b)
 ...the man that/who t has come
- (46) (a) ...den vent da Pol peinst da Marie t getrokken heet (= 3a)
 (b) *...den vent da Pol peinst die Marie t getrokken heet (= 3b)
 ...the man that Pol thinks that/whom Marie t has made a picture of

- (47) (a) ...den vent da Pol peinst da t gekomen is (= 4a)
 (b) ...den vent da Pol peinst die t gekomen is (= 4b)
 ...the man that Pol thinks that/who t has come

We shall take the cases of object relativization as a starting point (44 & 46). Their analysis is straightforward. The relative pronoun is moved from its argument position to COMP, and in (46) from COMP to COMP, into a position adjacent to the head of the relative construction. In that position it is deleted under identity with the head (cf. the derivation of (5) above). The same analysis applies to subject extraction in (45a) and (47a). The problematic cases are (45b) and (47b) which contrast with (44b) and (46b). If we take into account the analysis in the preceding section, the D-structure of a sentence like (45) will result in S-structure (48b), after application of *wh*-movement and cosuperscripting.

- (48) (a) ...den vent [_Sda [_Sdie gekomen is]] (D-str)
 (b) ...den vent [_Sdie_j daⁱ [_St_jⁱgekomen is]] (S-str)

In order to see how we might arrive at both (54a) and (45b) from (48b) we should take a closer look at the deletion under identity involved. The deletion affects the first element in a chain which ends in the trace in argument position. This chain is an $\bar{A}$ -chain, since it consists of one argument position, locally bound by an operator in $\bar{A}$ -position, which might be $\bar{A}$ -bound itself, and so on. This type of deletion under identity affects α_1 in an $\bar{A}$ -chain C, where $C = \alpha_1 \dots \alpha_i \dots \alpha_n$. An $\bar{A}$ -chain has, among others,¹⁵ the properties that (i) it has a θ -role and (ii) for each α_i , α_i binds α_{i+1} locally. The notion "binds" in this definition is defined as in (49):

- (49) α binds β iff α c-commands β
 and α and β are coindexed (by subscripts)

If we return to the cases of relatives with short extraction from subject position we may assume that there are two targets for deletion under identity: in (48b) there are two $\bar{A}$ -chains linking either *die* or *da* to the trace in subject position. They differ in that *die* is linked to the trace by movement and consequently by subscripting, while *da* is linked to the trace by agreement and thus by superscripts. Now we may explain the facts of (45) by extending relative deletion under identity in such a way that it comprises both cases. In order to do so, we extend the notion binding in the second condition on $\bar{A}$ -chains into BINDING and define BINDING as in (50).

- (50) α BINDS β iff α c-commands β
 and α and β are coindexed by sub- or superscripts

Now deletion of both *die* or of *da* in (48) will affect the first element in an $\bar{A}$ -chain with the appropriate properties. Pursuing the idea of obligatory deletion in COMP in West-Flemish relative clauses, we can explain the subject-object asymmetry, as indicated in (44)–(45). There is no superscript relation between AGR and an object, hence in the case of object relativization there is no $\bar{A}$ -chain between the inflected complementizer and the object position, and deletion under identity cannot affect the inflected complementizer (44b).

Finally we have to account for the cases of long extraction (46–47). The derivation of (47a) is given in (51).

- (51) (a) ...den vent [_S da [_S Pol peinst [_S da
[_S die gekomen is]]]] (D-str)
 (b) ...den vent [_S da [_S Pol peinst [_S die_j da
[_S t_j gekomen is]]]] (wh-mvt)
 (c) ...den vent [_S die_j da [_S Pol peinst
[_S t_j da [_S t_jⁱ gekomen is]]]] (wh-mvt)

In (51c) *die*, the relative pronoun is deleted under identity and the resulting sentence is (47a). How do we explain the grammaticality of (47b)? We will suggest that in (47b) it is not the relative pronoun *die* that is moved from the embedded COMP to the highest COMP, but rather the inflected complementizer *da*. We assume that *da* may be moved since it shares the relevant properties with relative pronouns in COMP: it contains pronominal features and it is related to an argument position in S (by superscripting).

Movement of *da* results in a structure like (52):

- (52) ...den vent [_S da_k da [_S Pol peinst
[_S die_j t_kⁱ t_jⁱ gekomen is]]]]

If we delete *da_k* under identity as the first element (α_1) in a chain (*da_k*, t_kⁱ, t_jⁱ) we arrive at the correct sentence (47b).

Again it is obvious that a similar derivation is not available in the case of object relativization: there is no chain from the complementizer *da* to the object position.

Summarizing our discussion of West-Flemish relative clauses, we have tried to show that the only assumption we need for an account of relativization is a rule of obligatory deletion in COMP under identity with the adjacent head of the relative construction. The subject/object asymmetries and the distribution of *die* and *da* follow from independently motivated assumptions and mechanisms, among which a theory of agreement as put forward in the preceding section. In our opinion the account provided here is preferable to an account which needs various unmotivated assumptions, as discussed in the first sections of this article.

Notes

1. See Haegeman (1983) for the formalization of such a rule
2. In fact Pesetsky relates this phenomenon to the Nominative Island Constraint (NIC). Given the fact that in the Government-Binding framework the explanation of these facts is taken over by ECP, we use ECP in stead of NIC. With respect to these phenomena the overall account remains the same.
3. For a more detailed discussion of these facts, see Haegeman (1983).
4. Dutch differs even more from West-Flemish than English or French, since in Dutch the lexical complementizer adjacent to the moved relative pronoun is obligatorily deleted. From the examples (1)–(4) only the ungrammatical (1b) and the grammatical (2b) correspond to the pattern of Dutch relatives. In Dutch there is no comparable subject/object asymmetry and consequently an analysis of West-Flemish along the lines of Dutch seems to be impossible.
5. It seems possible to derive at least part of this base structure from more general principles, among which the Extended Projection Principle. We shall come back to this and to the relevant literature on this topic (a.o. Den Besten (1982), Platzack (1982), Safir (1982)) in Bennis & Haegeman (forthcoming).
6. The position of INFL as the rightmost constituent of S is due to the fact that Dutch and West-Flemish are SOV-languages.
7. In what follows we shall take the latter option; for motivation and a more general account of this analysis we refer to Bennis & Haegeman (forthcoming). For the purpose of this paper it is sufficient to assume (24).
8. It is clear that the pronominal element *-se* which is part of the lexical complementizer, has to be distinguished from the pronoun *zie*. Without additional information one might think that *se* and *zie* together are one reduplicated form of the subject pronoun. However, in this type of sentence *-se* and *zie* may be separated by object clitics as in (i)

- (i) ...dase-t-ze zie gegeven heet
 ...that_[- plu, 3 pers, + fem] - it(DO) - her(IO) she(SUBJ) given has

This order is optional as is indicated in (ii)

- (ii) ...dase zie t-ze gegeven heet

Note that a kind of reduplication operation does indeed occur in West-Flemish in case of subject pronouns. This may lead to a sentence in which the subject pronoun is in fact three times visible, once as spelling out of the features of AGR, and two times in the reduplicated subject pronoun:

- (iii) ...dank ze kik gezien heen
 ...that_[- plu, 1 pers] her (DO) I (Redupl.SUBJ) seen have

9. See Den Besten (1982) for a discussion of these cases, which turn out to provide an argument in favour of his V_2 -analysis.
10. It is important to observe that we expect this complementizer agreement phenomenon only to occur in languages in which INFL is part of COMP. This expectation is borne out, since the only other language we know that shows these agreement phenomena is German, a regular Verb Second language like Dutch and West-Flemish. These German data are discussed in Bayer (this volume).
11. A third type of relation which is generally assumed to be indicated by superscripting is the relation between a pleonastic element and a phrase with which it is associated. Depending on the particular analysis of this type of construction, it is possible to extend the scope of definition (33) in such a way that these constructions are also covered by it.
12. Note that several other types of agreement phenomena fall under the definition (33).

NP-internal agreement, like the agreement between the noun and its determiner and/or adjective, may be conditioned by (33) under the definition of c-command as proposed by Sportiche and Aoun (1981). Secondly, subject-verb agreement can be seen as an instance of (33), since AGR does not only govern the subject but also the head of the V-projection, at least if we do not consider maximal projections like VP to be barriers for government.

Another extension of this definition might be to define agreement relations on categories of the type [+N] in stead of [+N, -V]. In that case we are able to extend (33) to agreement relations in predicative constructions with adjectives. Notice that such an approach enables us to make an interesting observation. We may divide the class of governors in two groups. Governors of type [-N], which are able to assign Case and governors of type [+N], which trigger agreement. So where the structural domain is defined in a uniform way by the definition of government, the type of operation that expresses relationships within this domain is dependent from the feature [+N]. Another consequence of this rather speculative idea is that INFL has an exceptional status. INFL is verbal, [-N], at least +Tense is verbal, and therefore INFL is able to assign nominative case to the subject. On the other hand INFL is nominal, [+N], at least AGR in INFL is nominal, and triggers agreement relations, as discussed in this article. It might be possible to derive subject/object asymmetries from this exceptional status of INFL.

13. With respect to (41) there may be another, alternative explanation. On the one hand the empty element is determined by a clitic-like element (*se*) in AGR, and is hence pro, i.e. pronominal. On the other hand it is also $\bar{A}$ -bound by the moved element in XP, hence a variable. This should thus mean that the empty position is an empty resumptive pronoun. But West-Flemish lacks the resumptive pronoun strategy, hence the sentence is ruled out.

14. One of the leading ideas in the discussion of the pro-drop parameter is the observation that the possibility of dropping the subject pronoun correlates with the possibility of inversion of the subject. In Bennis and Haegeman (forthcoming) we shall discuss the possibility of free inversion in West-Flemish. We shall try to show that West-Flemish does indeed allow free inversion of the subject, but that it is hard to make this inversion visible. The reason is that West-Flemish is an SOV-language. The effect of "Inversion" will be to adjoin the subject NP to the left side of the VP, the canonical position of objects. Since in non-inverted structures the subject is in a position directly to the left of VP too, it is difficult to show the difference. Evidence in favour of inversion is derived from the phenomenon of object clitics (cf. note 8) and the occurrence of an optional pleonastic element in front of the lexical subject as in (i) and (ii).

- (i) ...dan *tet* Pol en Jan *kommen*
 ...that_[+plu] it Pol and Jan arrive

- (ii) *Morgen komt tet* Jan
 Tomorrow comes it Jan

15. For a discussion about chains and their properties, we refer to Chomsky (1981). The exact definition of the notion $\bar{A}$ -chain is not very crucial for our purposes. What is important, is to extend the notion "bind" to "BIND"; a notation that is used in the same way by Chomsky (1981, p. 333) (cf. also Safir (1982b) though for a uniform indexing hypothesis).