

Conceptual Dependency and the Clausal Structure of Discourse

JOOST SCHILPEROORD AND ARIE VERHAGEN

Utrecht Institute of Linguistics OTS

1. Introduction

One prominent goal of discourse analysis is to uncover the conceptual structures denoted by a discourse. For many studies of discourse structure (Polanyi 1988, Mann & Thompson 1988, Sanders & Van Wijk 1996), analytical practice includes a characterization of the coherence structure of discourse, that is, the way in which the basic discourse units are connected to each other.¹ The notion *coherence structure* refers to connectedness of discourse that sets it apart from random sets of sentences. If discourse analysis is employed as a method for investigating cognitive processes in

¹ In this paper, we will confine ourselves to the analysis of monological, written discourse, i.e. *texts*.

reading and producing discourse, the coherence structure may reveal significant insights in the way people maintain or build a cognitive representation of texts (Van der Pool 1995, Sanders, Janssen, van der Pool, Schilperoord & van Wijk 1996, Sanders & Schilperoord *to appear*, Schilperoord 1996). However, a prerequisite for the analysis of discourse coherence is that the basic elements, or *discourse segments*, between which coherence relations hold, are identified by the analyst. An important question, then, becomes how one should proceed in dividing discourse into discourse segments. If, for example, a *Claim-Argument* relation is said to hold between two segments A and B, such that A expresses a claim and B an argument favouring that claim, then prior to this analysis and according to some criterion, A and B must have been labelled as 'discourse segments'. Such a criterion should meet various conditions. Apart from being feasible, it should properly identify the basic building blocks constituting a discourse. In addition, the criterion should be unbiased towards the theory of coherence relations underlying the analysis of discourse structure. Mann & Thompson (1988) have been particularly explicit on this matter: '(...) the division of the text into units should be based on some theory-neutral classification.'

The importance of this demand can be put into perspective by considering one outstanding definition of a coherence relation as *an aspect of meaning of two or more discourse segments which cannot be described in terms of the meaning of the segments in isolation* (Sanders, Spooren & Noordman 1992). This definition first of all presupposes a possibility to identify the 'segments in isolation'. Moreover, its validity crucially hinges on the availability of a procedure which does not already include this 'aspect of meaning' that cannot be described in terms of the isolated segments. For this reason, Mann and Thompson (and also Sanders & Van Wijk 1996) propose to employ a criterion based on the syntactic properties of sentences (the so-called *clause criterion*, to be discussed in more detail later on). In brief, they propose treating *grammatical clauses* as the basic discourse segments. Since clauses can, at least in theory, be identified on the basis of their structural properties, this criterion avoids inclusion of the meaning of segments, and the meaning of the relations holding between them, in identifying discourse segments. Mann and Thompson themselves have acknowledged that the criterion cannot be applied in a simple and straightforward manner. However, the exceptions they have therefore introduced are not, as we shall see, fully adequate and they furthermore lack explanatory import. It is these two problems that we want to address simultaneously in this paper.

Our first objective in this paper is to argue that a purely syntactic criterion sometimes makes predictions as to the segmentation of discourse that are counter-intuitive, and that where it does make plausible predictions it fails to *account* for this plausibility. Subsequently, we will show that a procedure for properly segmenting discourse should take into account the *conceptual* relations between clauses. We therefore introduce a notion of *conceptual dependency* between clauses, and show that where it holds between two clauses, these clauses cannot enter into a discourse-level coherence relation with each other. We propose a procedure for segmentation that takes this very notion as its point of departure, and we will show that it results in the right kind of segmentation, where the strictly syntactic procedure fails to do so. Section 2 highlights Mann & Thompson's proposal, and demonstrates a number of problematic cases of segmentation. Section 3 introduces and exemplifies the notion of conceptual dependency. The procedure that is based upon this notion is exemplified for several types of clauses. Section 4 works out some consequences of our proposal by relating it to the notion of 'elaboration sites', introduced by Langacker (1987).

2. Formally determined discourse segments

The clause criterion as it is applied by Mann and Thompson and by Sanders & van Wijk can be summarized as follows (see 1).

- (1)
 - i. Each clause is a segment, and a clause is a structure headed by a finite or infinite verb
 - ii. Exceptions to i are:
 - a. restrictive relative clauses
 - b. clausal subjects
 - c. clausal complements
 - iii. In cases of contracted coordinate clauses, the second conjunct is a segment provided only one major constituent is contracted.²

Now, consider a sentence like (2).

- (2) Because John refuses to eat he is visibly getting thinner

² We will ignore clause-complexes containing coordinate clauses in this paper.

There are two finite verbs, hence two clauses, in this sentence. The exceptions mentioned in rules ii and iii do not apply, so that the clause criterion splits the sentence into two segments: 'Because John refuses to eat' (= A) and 'he is visibly getting thinner'. The two segments in isolation can thus be identified as *John refuses to eat* and *he is visibly getting thinner*. The aspect of meaning that is not describable in terms of these segments is the causal relation between them, which, in the present case, is marked explicitly by the connector *because*. The coherence structure of (2) therefore is something like 'BECAUSE (segment a, segment b)'. The fact that *because* marks one segment as expressing a cause and the other as expressing a consequence cannot be derived from the meanings of segment a and b, for in isolation they simply designate two events. To sum up, for sentence (2) the clause criterion leads to a plausible analysis – the discourse segments are identified on the basis of their grammatical structure, and the coherence relation holding between them can be analyzed independently from the two segments. The causality relation is therefore situated on the level of *discourse*. Note that, in principle, it is not crucial for the causality relation to hold whether or not an explicit causality marker is present, or, if it is, what this marker is (cf. 'John refuses to eat; he is visibly getting thinner'). Sometimes the conjunction *and*, or punctuation, may suffice to allow the reader to construct the appropriate coherence relation between the two clauses/segments. In addition, the location of the coherence marker is of no relevance here either – compare for example 'John refuses to eat and *therefore* he is visibly getting thinner'.

However, there are some serious problems with the clause criterion. One of these shows up when we consider rule ii in (1). As this rule points out, several types of clauses are excluded from the set of possible discourse segments, namely *restrictive relative* clauses and *complement* clauses.³ Granting for the moment the definition of coherence relations, this implies that such clauses *cannot* enter into a coherence relation with another clause in the discourse. Therefore, the connections they entertain with their respective matrix structures cannot be located on the level of discourse but have to be located on the level of grammatical structure. The clause criterion thus makes a distinction between two types of clauses in terms of their ability to enter into a coherence relation (cf. Mann & Thompson 1988:248; Sanders 1992:115). This distinction is quite fundamental, but

³ We take clausal subjects to be a type of complement clauses (cf. below).

no motivation whatsoever is provided. Relative clauses and complement clauses are only listed as exceptions on rule i, without there being provided a proper (linguistic) justification for this exceptional status.

Another, related problem emerges in actual analytic practise. Consider (3), taken from a large corpus of Dutch formal judicial letters.⁴

(3) *Daarbij komt // dat zijn vrouw ernstig gehandicapt is // en dat hij een gezin heeft te onderhouden.*

(Thereby comes // that his wife severely disabled is // and that he a family has to take care of.)

To this it can be added that his wife is severely disabled and that he has to take care of his family.

Sentence (3) presents a highly frequent sentence types in the corpus. Especially the specific argumentative use of the verb 'komen' in 'Daarbij komt' ('To this it can be added') is highly frequent in Dutch formal prose. As to its partitioning, (3) consists of a main clause *Daarbij komt*, and two coordinated complement clauses *dat zijn vrouw ernstig gehandicapt is* ('that his wife is severely disabled') and *dat hij een gezin heeft te onderhouden* ('that he has a family to take care of'). Slashes indicate the clausal make-up of (3).

How many discourse segments are there in (3)? According to the clause-criterion (3) expresses *one* 'idea unit' because, according to rule (1)ii, complement clauses are considered part of their matrix constituent (regardless of *how many* complement clauses are governed by the same matrix). However, intuition suggests that (3) contains *two* segments instead of one for it expresses the ideas that 'his wife is severely disabled' and 'he has a family to take care of'. Between these two ideas the logical relation of addition holds (expressed by the conjunction *en* (and)) so that the coherence structure is something like 'AND (segment 1, segment 2)'. Clearly, this altogether plausible segmenting is prohibited by the clause

⁴ Much of the work reported here was motivated by the fact that we had to analyze the structure of formal judicial texts. Especially the division of these texts into discourse segments turned out to be problematic if the clause criterion was applied in the standard way (see Schilperoord 1996, 142ff.).

criterion. Consider another example, taken from the same corpus, containing a degree clause.

- (4) De situatie is echter in zoverre verschillend // dat in PLAATSNAAM reeds voor DATUM sprake was van accu-opslag (...)

(The situation is however to this extent different // that in PLACE already before DATE exists 'battery storage'.)

The situation differs to the extent that in PLACE battery storage had already existed before DATE.

Since the subordinate clause *dat in PLAATSNAAM reeds sprake was van accu-opslag* is neither a complement clause nor a relative clause, the clause criterion (1) evaluates (4) as containing *two* discourse segments, that is, the clause boundary corresponds to a boundary between two discourse segments. This analysis implies the existence of some coherence relation holding between the two clauses/segments. However, it's not easy to determine the nature of this relation. For instance, one cannot simply state this relation to be one of *specification*. To see the reason why, consider again the aforementioned definition of coherence relations. It explicitly preserves the meaning of the segments in isolation by defining coherence at a super segment-level. So, if the two clauses in (4) were two segments as well, then the following coherence relation can be deduced: SPECIFICATION (segment b, segment a). If this analysis were correct, then *omitting* segment b should only bear on the *coherence* relation between a and b, but *not* on the meaning of segment a. Omitting segment b would leave the meaning of segment b unspecified to be sure, but not entirely incomprehensible. However, deleting the subordinate clause results in a segment that *cannot* be properly comprehended (and one that is in fact plainly ungrammatical), as is demonstrated by (4a).

- (4) a ?De situatie is echter in zoverre verschillend

The comprehensibility/grammaticality of sentence (4a) can only be saved by deleting the adverbial phrase *in zoverre* as well. Actually, the tight connection between such degree expressions and degree clauses is also evidenced by the fact that various degree adverbs project specific complementizers (see Jackendoff 1977:202ff.), as testified by sentences (5).

- (5) Hij is *te* dronken *om* te staan (*He is too drunk to stand up*)
 Hij is rijk *genoeg om* een huis te kopen (*He is rich enough to buy a house*)
 Hij is *dusdanig groot dat* dit pak hem niet past (*He is so big that this suit doesn't fit him*)

These observations suggest that the connection between the two clauses in (4) is not one to be located at the discourse level. Again, however, this is obscured by the clause-criterion.

A third example of a frequent sentence type raising segmentation problems is (6).

- (6) Te uwer informatie merk ik nog op dat cliënt voorziet dat het niet eenvoudig zal zijn om snel ander werk te vinden.
 (For your information remark I also that client foresees that it not simple shall be to soon other work to find.)
For your information I should also like to add that my client foresees that it will not be simple to find another job in the short term.

Sentence (6) presents a case of deep embedding of the clause expressing the actual substance of the compound sentence. The clausal make-up of (6) is presented in (7).

- (7) Te uwer informatie merk ik nog op //
 dat cliënt voorziet //
 dat het niet eenvoudig zal zijn //
 om snel ander werk te vinden.

Verbs participating in the first two clauses express statements of communicative intent (*opmerken*, 'inform'), and of belief (*voorzien*, 'foresee'). Such verbs typically require a complement clause for their conceptual completion, so here there are no problems in applying rule (1). Things are different, however, with the deepest embedded clause 'om snel

ander werk te vinden'. As this clause seems to elaborate the adverbial element 'niet eenvoudig' ('not simple'), rule (1) draws a segmental boundary between the first three composite clauses and the fourth one. Again, intuition tells us such a decision is off the mark. Stated rather loosely, it seems that the first two clauses express some kind of *perspective on the assertion expressed in the final clause*, whereas the third clause *evaluates that assertion*. To put this in more appropriate theoretical terms, the first two clauses may be said to denote a number of connected *mental spaces* of which the content is ultimately designated by the final clause (see Fauconnier 1985, Sanders 1994). Therefore, an adequate procedure for classification should do justice to this kind of dependency and evaluate (6) to be *one* discourse segment.

We could, of course, proceed by presenting alternative types of compound sentences raising segmentation problems. However, we believe that the problematic cases discussed so far exemplify a more general problem with the clause criterion – the fact that it ignores the *conceptual* relations that may exist between different types of clauses. Whether or not a *syntactic* clause is a discourse segment as well crucially depends on how the information expressed by this clause relates conceptually to the information expressed by other clauses within the structure. We want to propose a method for segmenting which is based on a notion that we will refer to as *conceptual dependency*. In brief, one clause is conceptually dependent upon another clause, if its semantics cannot be conceptualized without essential reference to the conceptualization of another clause. If such an interdependency exists, we argue that the two clauses cannot enter into a coherence relation with each other. Moreover, as will be demonstrated, *conceptual dependency* not only preserves the segmentation decisions that follow from the original clause criterion, but additionally accounts for the rather arbitrarily listed exceptions concerning relative clauses and clausal complements.

3. Conceptual dependency

If defined at the proper level of abstraction, the problems encountered in the segmentation of sentences into component clauses that correspond to discourse segments can be solved as soon as we acknowledge the kind of *conceptual dependency* relations holding between clauses. This notion was already hinted at in the former section, here we will elaborate it in more detail. We shall do so by discussing a classic grammatical problem: the distinction between restrictive and non-restrictive relative clauses.

In Verhagen (1996), an analysis is presented of different types of relative clauses independently of the problem of discourse segmentation,

which will allow us to make a first move towards a better understanding of the relationship between clause-status and segment-status. Consider example (8).

- (8) I liked the man who took that dog with him even more than the one who brought it in.

Here the conceptualization of (the intended referent of) *the man*, in this particular sentence, can be said to be **dependent** on the conceptualization of the relative clause. Without the latter information, the conceptualization of the object of "like even more" is incomplete. So what we can say is that a restrictive relative clause is one that the nominal head is conceptually dependent on. In contrast, consider the nominal head in (9):

- (9) They asked the vicar, who so far has refused to state his personal opinion, to act as an intermediary.

In this case, the conceptualization of the object of "ask" is not dependent on the information in the relative clause; so we may say that a relative clause is non-restrictive if its head is **not** conceptually dependent on the contents of the clause. Verhagen therefore suggested that the **interpretive** relations between both types of relative clauses and their matrix clauses are different. Following this suggestion, we can say the following:

- A restrictive relative clause is a part of the conceptualization of a participant in the event denoted by the matrix clause, hence an integral part of the conceptualization of the matrix;
- A non-restrictive relative clause is not an integral part of the conceptualized event of the matrix; it has a much more 'loose' conceptual connection to the matrix; for instance, Daalder (1989:202) calls the function of such clauses "adverbial-like" (we will get back to this parallel later).

Now it is only natural to say that a unit that is an integral part of the conceptualization of a clause cannot enter into a coherence relation with that clause on the level of discourse interpretation. Given that it is **conceptually integrated** into the higher clause, it is no longer available as a separate discourse segment, as it were. Now if the essence of a restrictive relative clause is that some part of the **matrix** is conceptually

dependent on it, as we just suggested, we may formulate the following hypothesis on the relation between such conceptual dependency and the possibility of a status as a separate discourse segment:

If a constituent of clause A is conceptually dependent on a clause B, B is an integral part of the conceptualization of A, and therefore not available as a separate discourse segment (cannot enter into a discourse coherence relation with A, or any other part of the discourse).

For the discourse analyst this means that a matrix element being conceptually dependent on a subordinate clause, makes it impossible to treat the left hand boundary of the subordinate clause involved as the beginning of a separate discourse segment.

Now that we have formulated the idea in this general way, how does it apply to the problematic cases that we introduced earlier? Take the frequent type (3), repeated here as (10).

- (10) Daarbij komt dat zijn vrouw ernstig gehandicapt is en dat hij een gezin heeft te onderhouden

In fact, the idea applies straightforwardly. The phrase *Daarbij komt (dat)* in itself is **necessarily** incomplete; speakers of the language can be said to **know** that it is incomplete as an element of use: for the phrase to make sense in actual usage, it is indeed dependent on the contents of the actual subordinate clause. So there can be no discourse segment starting at the first occurrence of *dat*. However, when we have reached the end of the first subordinate clause (...*is*), we have a situation where the conceptualization may be completed, so it is not dependent on what follows; hence we do not only have a new clause starting here, but also a new discourse segment. In this way, then, this string contains two discourse segments, which is precisely what we wanted: intuitively, there are **two** coordinated additional supporting arguments to what has already been stated before.

This is not to say that restrictive relative clauses and this type of complements (actually: a subject clause) are similar in **all** respects. In one sense, they are even in contrast: a restrictive relative does not constitute a discourse segment, whereas the subordinate clause in (10) actually provides the substance of the relevant discourse segment. However, this is

completely due to the difference between the conceptual roles of the **matrix constituents** that the respective clauses are subordinated to. As we said, in restrictive relatives, this phrase denotes a participant in the event depicted by the matrix clause. The matrix phrase (in this case the entire matrix clause) in (10) does not even relate to the content level of the discourse, but rather to the **speech act level**.⁵ That is, the element *Daarbij* refers not to the **contents** of preceding discourse, but instead to the *act of uttering this information, argumentation, or whatever speech act the speaker/writer performed with the previous discourse segment*. Consequently, the conceptual contents of the syntactic main clause in (11) is (at least) one level 'higher' than the contents of the preceding main clause; graphically:

(11)

<i>speech act level:</i>	Daarbij komt dat	en dat
<i>content level:</i>	preceding	zijn vrouw ernstig	hij een gezin heeft te
	argumentation/	gehandicapt is	onderhouden
	information		

By the same token, the contents of the two **subordinate** clauses with *Daarbij komt dat* are conceptually on the same level, as the figure indicates, as the **main** clause preceding them. In other words: If the speech act part and the content part are distributed over a matrix and a subordinate clause, these two parts do not constitute separate discourse segments, but this does not make the subordinate clause less a subordinate one (i.e. as long as the main clause is perceived as such), namely subordinated to the explicit expression of the speech act level part. Thus we want to claim that applying these independently motivated distinctions to the problem at hand, allows one to say **at the same time** that the contents of some subordinate clauses is in fact not subordinate (i.e. at the content level in relation to other discourse segments), and is (i.e. to the particular speech act in which it is being produced).

⁵ Loosely speaking, our main point is that it relates to a higher level than the content level. Perhaps we should just call it 'metalinguistic', but nothing in our argument here hinges on the choice of terminology.

The situation of a restrictive relative clause may, in contrast, be depicted graphically as follows:

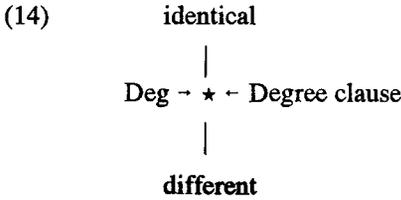
- (12) *content level:* I liked the man even more than the one who
 took that dog with him who brought it in

Here, the fact that there are syntactic subordinate clauses does not result in a difference of the level at which the conceptual contents is relevant, since the heads themselves are already functional at the content level. But this does not alter the fact that we are allowed to make the generalization proposed above: Whenever a subordinate clause provides information that (part of) the matrix is conceptually dependent on, this clause does not constitute a separate discourse segment.

Let us finally have a look at conceptual dependency relations in sentences containing a complex consisting of a degree expression and a degree clause. Consider (4) again, repeated here as (13).

- (13) De situatie is echter in zoverre verschillend dat in PLAATSNAAM
 reeds voor DATUM sprake was van accu-opslag (...)

The structural regularities in degree clauses captured in section 2 can be subsumed under the header of conceptual dependency. Consider the semantics of the *Deg*-phrase ‘in zoverre’. This phrase evokes the idea of some kind of scale, of which the poles are defined by the adverbial phrasal head ‘verschillend’. The scale thus ranges from ‘completely identical’ to ‘completely different’. However, the *Deg* element not only evokes the scale, it also suggests a particular point on that scale. The tight structural connection it entertains with the degree clause can then be accounted for conceptually if it is acknowledged that the scale itself is ‘calibrated’, as it were, by the subordinate degree clause ‘dat in PLAATSNAAM...’. And since one cannot conceptualize the scale without some means of calibration, there exists conceptual dependency between the two clauses. This rather complex semantic structure is captured schematically by diagram (14).



Together, the conceptualization has an argumentative stance. The dependent clause *justifies* the scale point indicated by the adverb 'verschillend', but it only does so because of the designated scale itself. The idea that a clause such as (15):

(15) ?De situatie is in zoverre verschillend

is conceptually 'incomplete', so to speak, arises from the insertion of 'in zoverre'. For example, it is very well possible to state:

(16) De situatie is verschillend / De situatie is anders

Hence, it can be concluded that a dependency relation holds between the dependent clause and the element 'in zoverre'.

Returning to our main theme now, we want to suggest the following extension of our proposal. As we have seen, a relative clause in itself allows for a two-way interpretation: It either is restrictive, thus not a separate discourse segment, or it is non-restrictive and has the status of a segment. However, other types of syntactically subordinate clauses do not allow such a choice, but have to be taken as either one type or another, independently of the context. Adverbial clauses, introduced by different kinds of conjunctions, never provide information that some matrix part is conceptually dependent on, thus they always constitute discourse segments. On the other hand, the (first) complement of a predicate (both in subject- and in object-functions), introduced by the non-specific complementizer *that*, always specifies information that the matrix is conceptually dependent on, and thus never constitutes a discourse segment. Then again, a second (third, etc.) complement added on to (by means of coordination) to the first one, does not necessarily provide

information that the matrix is conceptually dependent upon, and thus may be a discourse segment on its own.⁶

4. E-sites

In the final section we want to link our proposal to Langacker's notion of 'elaboration-sites' (e-sites) (Langacker 1987, 1991). Throughout this paper we have been using the notion 'dependence' in a way that is not standard in general linguistic practice: Usually all syntactically subordinate constituents are considered 'dependent', while we have applied this notion to matrix phrases the conceptualization of which is, in some sense, not 'complete' without the embedded information. This actually runs parallel to Langacker's notion of conceptual dependence:

D is conceptually dependent on A to the extent that A elaborates a salient substructure of D. (Langacker 1991:436)

In the case of relational predicates, this 'salient substructure' is a part of the predicate's representation, called an *elaboration site* or *e-site* (Langacker 1987:304); in non-relational predicates it is something constructed on a more ad hoc basis; we will suggest, however, that conceptual dependence is neutral as to the source of the relevant substructure.

An e-site of a particular structure is a more or less *fixed* schematic component of that structure which is subsequently elaborated by another structure. An *elaboration*, then, involves an instantiation of an e-site which is consistent with its schematic specification, but which is also 'more fully and precisely specified' (Langacker 1987:489). The relation between the two structures⁷ can then be quasi-formalized as in (17) in which the subscript *De* denotes the e-site of D.

(17) D (_{De} A)

Although Langacker did not envisage the types of phrases and relations that we are concerned with, and, more specifically, although his proposal

⁶ Note that we are only concerned here with the determination of segment-status, not with the place of a segment in the (hierarchical) text structure.

⁷ Compare Langacker's treatment of *valence relations* (1987, 277ff.).

did not involve the problem of discourse segmentation, it is not difficult to extend it to those areas.

The characteristic function of complement clauses is precisely to elaborate a salient substructure of the main clause: the e-site of the main verb, i.e. a relational predicate. These substructures can be assumed to be 'salient' as they refer to prominent participants in the event denoted by the main clause. In terms of (17), these interrelations can be stated as follows (18).

(18) (...V (V_{e-})_{comp.cl.}...)main cl.

Because complement clauses elaborate the e-site of the main clause verb, their conceptualizations are an integral part of the conceptualization of the main clause. Without these, the conceptualization of the main clause would be necessarily incomplete, which is the reason why in these cases main clauses are conceptually dependent upon subordinate clauses. It then follows from Sanders' (1992) definition of coherence relations that complement clauses cannot enter into a discourse relation with the main clause. After all, this definition *presupposes* the possibility of independent conceptualizations of the clauses/segments between which a coherence relation holds.

The difference between these types of clauses and restrictive relatives can be described, in a way we will now explain, in terms of a difference of entrenchment of the relationship between main clause elements and their respective e-sites. Typically, illocutionary verbs, verbs of belief and the argumentative use of the verb 'komen' in formal Dutch prose are strongly associated with a schematic e-site. One may say that structures such as V (V_{e-})_{cc} are part of the speaker's knowledge of his language, or his *Mental Grammar*: knowing a verb such as *note* or *propose*, or a typical fixed phrase such as *Daarbij komt* (cf. "It should be added"), simply involves knowing that it requires complementation in one way or another as well. Thus the salience of these substructures (e-sites) is a matter of memory, not dependent on particular conditions of use.

In the case of nominals, which are not even relational predicates to begin with, things are different. What we have to be crucially aware of here is that when put to use, nominals are not so much used to evoke the concepts associated with them in memory, but to evoke, *via* these concepts so to speak, the conceptualization of a particular ('grounded') *instance* of

that concept.⁸ From a communicative point of view, it is thus pivotal that a speaker allows his interlocutor to establish 'mental contact' with that particular instance of a concept. In other words, the speaker must provide his interlocutor with means to get access to the intended noun referent (cf. Langacker 1990:321). The notion of *accessibility* of noun phrase referents is now well entrenched in the cognitive linguistic literature (see especially Ariel 1988, 1990). Languages contain rich systems of linguistic means to make intended nominal referents accessible for listeners, and to signal the current accessibility status of a nominal referent. For example, (unstressed) pronouns, or zero-anaphors ('handle with care') signal high accessibility of the nominal referent, whereas full-fledged NP's containing adjuncts usually signal low accessibility. In addition, determiners and quantifiers are used to code accessibility: The instantiation of a new referent within a discourse is typically established by using indefinite articles.

What we want to propose here is that restrictive relative clauses can be treated as a way for the speaker to signal the level of accessibility of a nominal referent. They literally *restrict* the selection of instances from a set denoted by a concept. That is, a restrictive relative clause does not elaborate an e-site that is given by linguistic knowledge, but it does elaborate a conceptual substructure of the intended nominal conceptualization. One might suggest that elaborating a conceptual substructure of an intended nominal conceptualization is thus conceptually parallel to the elaboration of an e-site of a relational predicate – only here it is one that is situationally evoked, determined by the accessibility of the intended discourse referent. Arguably, if the intended referent of a noun is relatively *new* in a discourse, or if it has to be reestablished, or if it is to be distinguished from referents competing for attention with the intended one, the function of restrictive relative clauses may be one of identifying its referential domain. In such cases the relative clause indeed does elaborate a situationally evoked salient substructure of the *intended conceptualization* of the main clause. This situation contrasts with the one where a noun is accompanied by a nonrestrictive relative clause. Consider the sentences in (20):

- (20) a. I liked the man *who took that dog with him*,....
 b. I liked the man, *who took that dog with him*,....

⁸ In principle, the distinction is relevant in the case of verbs as well, but it is not so crucial for the sense of conceptual dependence that we are concerned with here.

Langacker (1991:431) argues that the head noun elaborates a salient participant of the relative clause, and in that sense one might say that the relative clause is dependent on its head noun; this holds for both (20)a and b, as the trajector of the relative clause in each is only schematically indicated by the relative pronoun, indicating a participant elaborated as *the man* in the matrix clause. However, in sentence (20)a there is an additional dependency in that the denoted referent of the head noun is made accessible by the relative clause (see also Deane 1992:105, who describes this dependency in terms of 'referential dependency'). In other words, it is only by virtue of the information provided in the restrictive relative clause, that the intended referent of the noun-phrase can be established, or made accessible. It is in that sense, therefore, that we propose that the conceptualization of the noun phrase is dependent upon the restrictive relative clause. Once again, without the restrictive relative clause, the conceptualization of the head-noun by itself does not suffice. For that reason, the restrictive relative clause cannot be regarded as a distinct discourse segments. In (20)b, on the other hand, the relative clause only serves to modify, and not to *identify* the intended conceptualization of the noun phrase; in that case, the head noun does suffice for identification, and therefore the conceptualization is *not* dependent upon the relative clause. For that reason, non-restrictive relatives *are* distinct discourse segments that enter into a coherence relation with the main clause (a relation that can be, for instance, 'specification' or 'modification', and sometimes even concession; cf. Daalder 1989).

The difference between, on the one hand mental state verbs or illocutionary verbs, which contain an e-site to be elaborated by a complement clause, and nouns on the other, is thus a matter of degree of fixedness. Whereas the verbs mentioned here typically contain fixed e-sites of the structure $V (v_e^-)_{cc}$, the strength of the connection between a nominal concept and the elaboration of a substructure of that concept is much more loose, and dependent upon the current discourse situation. Once the referent of a noun phrase is established properly, the need to identify that referent vanishes, at least for some time during discourse processing. That is, if nouns refer to concepts already *given*, or inferrable from the previous discourse, relative clauses cannot be considered to elaborate a substructure of the nominal concept (and are therefore *non-restrictive*).

Now, let us have a look at clause complexes containing connecting elements, such as *because* or *while*. Such connectors are prototypical

means for indicating relations between clauses on the discourse level. In Langacker's view, they even go so far as to *profile* the interclausal relationship (ibid, 426). *While*, for instance, results in temporal inclusion of the events denoted by two clauses. As this typically requires the contents of these clauses to be conceptualized independently, *while* defines clausal connections on the discourse level, which is perfectly in line with the definition of coherence relations. In Langacker's analysis, *while* contains *two* e-sites, called the *trajector* and the *landmark* respectively, which are elaborated by the two connected clauses (i.e. the main clause and the subordinate clause). It follows that a subordinate clause elaborates the e-site of a clause connector C, *not* of a clause participant, which is the very reason to consider these (adverbial) clauses as discourse segments. This is schematized in (21).

$$(21) \quad C ((C_{e1} \cdot \cdot)_{c1}, (C_{e2} \cdot \cdot)_{c2})$$

The independent nature of the clauses 1 and 2 is accounted for by not letting one clause depend on the other one, whereas the autonomous status of the connector C is accounted for by not letting it participate within one of the connected clauses. In a sentence such as (22) (taken from Langacker 1991), the contents of the clauses can be conceptualized independently, but their respective contents are modified by 'placing' them together in some specific context.

$$(22) \quad \text{While} ((\text{she was working in the garden})(\text{Janet found a lizard}))$$

Rather than going into the details with respect to the impact the contents of these clauses have upon each other (see Langacker 1991:425), we maintain that, generally speaking, these effects result from a specific discourse relation which is imposed, so to speak, on the structure as a whole. We emphasize this point because an adverbial clause might have the effect of conceptually modifying the contents of a main clause (which, then, would be a type of conceptual dependency). Therefore we conclude this paper by discussing examples such as (23), taken from Pander Maat (1994).

$$(23) \quad \text{Je werkt tot de bel gaat} (\text{You work till the bell goes, } \textit{You must work till the bell rings})$$

Pander Maat argues that (23) expresses *one* discourse segment, and not *two* as we would say, because the subordinate clause has the effect of *restricting* the meaning of the main clause (cf. Pander Maat 1994:33, see also Haliday 1987). Sentences such as (23) thus seem to run counter to the distinctions made here. So how are we deal with it?

We think we simply have to be explicit about the possible relations between conceptual (in)dependency of syntactically subordinate clauses and the elements introducing these clauses (phrases introducing the relevant clauses and the words in the 'complementizer'-position). In principle, we have three logically possible kinds of relations:

- a) if a clause C is introduced by element X, the matrix is always conceptually dependent on C (X marks C and its matrix as conceptually dependent);
- b) if a clause C is introduced by element X, the matrix is never conceptually dependent on C (X marks C and its matrix as conceptually independent);
- c) the matrix of a clause C introduced by X sometimes is conceptually dependent on C, and sometimes it is not (X in itself does not mark C and its matrix as conceptually dependent or independent).

Relative clauses are clearly of type c): the 'simple' fact that a clause is introduced by a relative pronoun is in itself insufficient to decide whether the head noun is conceptually dependent on the clause, or if the clause constitutes a separate discourse unit. In such cases, conceptual (in)dependency is purely a matter of the meaning of the entire structure.

The cases that constituted the problem that we started out with are of type a): Whenever we have a phrase like *Daar komt bij dat*, we know this to be conceptually dependent on the clause following it, so that the latter is not a separate discourse unit. In fact, we would ultimately like to explore the possibility that the conjunction *dat* ('that'), the marking typical of complement clauses, functions precisely as a lexical manifestation of a relation of conceptual dependency between clauses.

Adverbial clauses like (20), introduced by conjunctions such as *omdat* ('because'), *terwijl* ('while'), etcetera, are of type b): Whenever we encounter *because*, *while*, etc., we know that the clauses surrounding or

following it are separate discourse units, and that its matrix is not conceptually dependent on these clauses.⁹

We would like to suggest now that a subordinate clause introduced by the preposition *tot* is of type c), just like a relative clause: being introduced by *tot* (like being introduced by a relative pronoun) is in itself a strict marking neither for conceptual dependency, nor for independence. In (21), the meaning of the entire structure apparently imposes a relation of conceptual dependency (the subordinate clause is taken as a restrictive modifier, in this case of the predicate), at least as a kind of default. Such relations in fact parallel the use of *tot* as a straightforward preposition with a nominal complement, as in *Je werkt tot zes uur*, 'You will work till six o'clock'. In other cases, however, *tot* introduces a non-restrictive clause; an example is (24):

- (24) Jan heeft aan het slothoofdstuk van zijn dissertatie zitten werken, tot(dat) de bel het diner aankondigde.

(John has on the final-chapter of his dissertation been working, until (that) the bell the dinner) announced

John has been working on the final chapter of his dissertation, until the bell announced dinner.

It may be the case that in such cases the more elaborate alternative *totdat* ('compound' of *tot* and the subordinating conjunction) is in fact preferred over the simple *tot* - something that would be expected from the point of view of iconicity. However, intuitions are subtle in this area, and we have not yet undertaken a further investigation of this suggestion. Be this as it may, one can in principle 'force' a non-restrictive interpretation on the *tot*-clause in (23) (indicated in writing by punctuation), without *tot* becoming impossible - indicating that *tot* in itself is compatible with both types of interpretations. In any case, this would mean that there is a class of

⁹ Note that this idea, if correct, might provide a basis for explaining the intuition that *dat* is semantically more 'empty' than other subordinating conjunctions. That is, we would still maintain that *dat* has a meaning, i.e. has a cognitive function to fulfill, but that its specific function is purely 'relational': an operation on the relation between the conceptual contents of two clauses (dependency), whereas *because*, *although*, etc. primarily denote a relation of a certain kind (causal, concessive, etc.), and as a consequence also indicate a relational operation (in this case: of conceptual independence). We hope to explore these ideas further in later work.

expressions introducing subordinate clauses, which does not determine the profile elaborated by the clause (specifically: *dat*), so that such a clause is always conceptually integrated into its matrix. On the other hand, there is a class of expressions which does determine the subordinate clause's profile; such clauses may or may not be conceptually integrated: certain elements necessarily mark their clauses as separate discourse units, while others (like *tot*) are indeterminate in this respect.

Our point of departure in this paper was Mann and Thompson's conviction that semantic factors should not interfere with the division of discourse into discourse units. We have attempted to argue that their formal criterion does not always arrive at a proper segmentation. Also, we have argued that problematic cases for a formal criterion always, in one way or another, involve conceptual dependency of clauses. By doing so, we have offered a semantic account for the list of exceptions on the formal clause criterion mentioned by Mann and Thompson. Restrictive relative clauses, and complement clauses cannot be distinct segments because they elaborate a salient substructure within their matrix. In addition, the notion of conceptual dependency has been extended to other types of clause-complexes, notably complexes involving degree-clauses. To conclude, although in its present form our proposal applies to a restricted set of 'problematic' clausal complexes, we claim that its implication exceeds this domain. Needless to say, however, more research will be necessary in order to arrive at a conceptually based, sound criterion for segmenting discourse. We maintain, however, to have made a case for such a conceptual approach to this problem by demonstrating that applying insights from cognitive linguistics is crucial to the very basis of discourse analysis: the identification of relevant discourse units.

Acknowledgements

We would like to express our gratitude to Henk Pander Maat, Ted Sanders, and an anonymous reviewer for their comments and advice.

Authors' address: Utrecht institute of Linguistics OTS, Trans 10, 3512 JK Utrecht, The Netherlands. Email: Joost.Schilperoord@let.ruu.nl, Arie.Verhagen@let.ruu.nl.

References

- Ariel, M. 1988. Referring and accessibility. *Journal of Linguistics* 24, 65-87.
 Ariel, M. 1990. *Accessing noun-phrase antecedents*. London: Routledge.

- Daalder, S., 1989. Continuative Relative Clauses, in N. Reiter, ed., *Sprechen und hören. Akten des 23. Linguistischen Kolloquiums*. Tübingen: Niemeyer.
- Deane, P.D., 1992. *Grammar in Mind and Brain. Explorations in Cognitive Syntax*. Berlin: Mouton de Gruyter.
- Fauconnier, G., 1985. *Mental Spaces: Aspects of Meaning Construction in Natural Language*. Cambridge, Mass.: MIT Press.
- Jackendoff, R., 1977. *X-bar syntax: a study of phrase structure*. Cambridge Mass.: MIT Press.
- Langacker, R.W., 1987. *Foundations of Cognitive Grammar, Volume I Theoretical Prerequisites*. Stanford: Stanford University Press.
- Langacker, R.W. 1990. *Concept, Image, and Symbol. The Cognitive Basis of Grammar*. Berlin: Mouton de Gruyter.
- Langacker, R.W., 1991. *Foundations of Cognitive Grammar, Volume II Descriptive Application*. Stanford: Stanford University Press.
- Mann, W.C. & S.A. Thompson, 1988. Rhetorical Structure Theory: toward a Functional Theory of Text Organization. *Text* 8, 243-281.
- Pander Maat, H., 1994. *Tekstanalyse, een pragmatische benadering*. [Text analysis, a pragmatic approach.] Groningen: Martinus Nijhoff.
- Polanyi, L., 1988. A formal model of the structure of discourse. *Journal of Pragmatics* 12, 601-138.
- Pool, E. van der, 1995. *Writing as a conceptual process. A text-analytical study of developmental aspects*. Doctoral dissertation, University of Tilburg.
- Sanders, J., 1994. *Perspective in Narrative Discourse*. Doctoral Dissertation, Tilburg University.
- Sanders, T., 1992. *Discourse structure and coherence; Aspects of a cognitive theory of discourse representation*. Doctoral dissertation, University of Tilburg.
- Sanders, T., W. Spooren & L. Noordman, 1992. Toward a taxonomy of coherence relations. *Discourse Processes* 15, 1-35.
- Sanders, T., D. Janssen, E. van der Pool, J. Schilperoord & C. van Wijk, 1995. Hierarchical structures in writing products and writing processes, in G. Rijlaarsdam, H. van den Bergh & M. Couzijn, eds., *Theories, models and methodology. Current trends in research on writing*. Amsterdam: UvA-press.

- Sanders, T. & C. van Wijk, 1996. PISA - A Procedure for Analyzing the Structure of Explanatory Texts. *Text* 16, 91-132.
- Schilperoord, J., 1996. *It's about Time; Temporal Aspects of Cognitive Processes in Text Production*. Amsterdam: Rodopi.
- Schilperoord, J. & T. Sanders, *Pauses, Cognitive Rhythms and Discourse Structure*, to appear.
- Verhagen, A., 1996. Sequential Conceptualization and Linear Order, in E.H. Casad, ed., *Cognitive Linguistics in the Redwoods; the Expansion of a New Paradigm in Linguistics*. Berlin/New York: Mouton de Gruyter.