

Degand, Liesbeth & Sanders, Ted (1999) Causal connectives in language use. Theoretical and methodological aspects of the classification of coherence relations and connectives, Levels of Representation in Discourse, *Working Notes of the International Workshop on Text Representation*, Edinburgh University, 3-12.

Causal connectives in language use. Theoretical and methodological aspects of the classification of coherence relations and connectives

Liesbeth Degand, University of Louvain (Louvain-la-Neuve)
Ted Sanders, Utrecht University

1. Classification of coherence relations

In recent years, a significant part of research on coherence relations has focused on the categorization of coherence relations (CR), trying to answer the question: how can the many different sets of relations be organized? A dominant distinction in several, if not all classifications, is that between so-called content, ideational, external or semantic relations on the one hand, and presentational, internal and pragmatic relations, on the other hand. In our view, criteria like these, in terms of which sets of relations can be ordered, can be considered as LEVELS OF REPRESENTATION of coherence, that is, relations can be understood to hold on a propositional level, or on an illocutionary level.

This systematic difference between types of relations has been noted by many scholars working on discourse coherence. Still, there is a lot of discussion about the exact definition of such distinctions (see e.g. Bateman & Rondhuis, 1997; Degand, 1996; Hovy, 1990; Knott & Dale, 1994; Knott, 1996; Martin, 1992; Moore & Pollack, 1992; Pander Maat, 1998; Sanders, 1997; Sanders & Spooren, 1999). At the same time, when we take a look at the examples researchers give to illustrate their distinctions, it appears that they agree on the characteristics of the prototypical relations (Sanders, 1997), in fact they agree in about 80% of the cases. For instance, the clear cases of speech-act / pragmatic / presentational / internal relations are the ones in which it is impossible to construct a relation at the propositional level, as in the (indirect) speech act example in (1).

(1) I'm busy. You can take your own beer out of the fridge.

Prototypical epistemic / pragmatic / presentational / internal relations are cases in which the writer argues for something she claims to be true. Most of the times, the writer is clearly present as the communicating person, and the unambiguous cases are those in which it is impossible to get a semantic / subject matter / internal / content reading, such as Sweetser's example given in (2). They often contain linguistic elements expressing the evaluation from the perspective of the author.

(2) Since John isn't here, he has evidently gone home.

Prototypical content / semantic / subject matter / external relations concern events which have already taken place, such as Sweetser's (3), so that there can be no dispute about the "truth" of the statement.

(3) Since John wasn't there, we decided to leave a note for him.

2. Classification of connectives

It is particularly interesting that some of the classifications explicitly concern relations, whereas others concern connectives. Sweetser (1990), for instance, argues that the ambiguity in the use of conjunctions can be explained in terms of three domains. Her famous examples (4)-(6) illustrate that English *because* has three readings: a content reading, an epistemic reading and a speech act reading.

Degand, Liesbeth & Sanders, Ted (1999) Causal connectives in language use. Theoretical and methodological aspects of the classification of coherence relations and connectives, Levels of Representation in Discourse, *Working Notes of the International Workshop on Text Representation*, Edinburgh University, 3-12.

- (4) John came back because he loved her.
- (5) John loved her, because he came back.
- (6) What are you doing tonight, because there's a good movie on.

It is important to note this similarity in classification of CRs and connectives. One may expect that a valid classification of CRs, which are considered as cognitive entities, has some counterpart in language. After all, there are restrictions on the use of connectives. And there are many correspondences between connectives and the 'relational' features they signal. Even though there is no one-to-one mapping between relations and connectives, the restrictions on the use of connectives imply an organization of the relations they can express; they do not just co-exist as a set of relations on one and the same level.

A strong form of this idea is that connectives exist which can only be used to express one type of coherence relation, or to put it in Sweetser's terms, are domain-specific. Given the supposed 'basicness' of conceptual distinctions like *semantic / content* vs. *epistemic / pragmatic*, it is attractive to assume that these distinctions are similar to the restrictions on the use of connectives. In such a view, the lexicon of connectives would reflect the basic distinctions between coherence relations in a systematic way.

Now, no matter how attractive the parallels and no matter how suggestive the similarities between relations and the linguistic devices that express them, the categorization of these linguistic devices is as yet hardly based on data of their distribution in actual language use. In fact, none of the ideas on 'domain-specificity' of connectives and cue phrases are based on corpora. By contrast, recent corpus studies do reveal that existing categorization proposals *cannot* account for the data of connective distribution in a straightforward way (Degand, 1996; Pander Maat, 1998; Pander Maat & Sanders, 1995).

3. Toward an empirical approach

In our view, current suggestions on the 'domain-specificity' of connectives should be considered as hypotheses to be tested in empirical research of language in use. This is exactly the purpose of the research we want to address here: To study the distribution and use of connectives and investigate to what extent their distribution can be described in terms of the existing categorizations.

We look upon the lack of an empirical basis as a serious obstruction for the further development of the domain theory of connectives, and in this paper we want to explore the possibilities to improve this situation. Which are the available methods for empirical testing? What implications do they have for the question of categorization of the connectives? And can they be used in such a way that they enable us to compare over different contexts, different analysts and different languages? For the sake of clarity, we focus on causal connectives.

3.1. The substitution test and the paraphrase test

Currently, two methods are most used to study the meaning of connectives in corpora: the substitution test and the paraphrase test.

In a substitution test, a given connective is substituted by another one (cf. Knott & Dale, 1994; Knott, 1996 for a systematic study of English cue phrases and Knott & Sanders, 1998 for a contrastive study

Degand, Liesbeth & Sanders, Ted (1999) Causal connectives in language use. Theoretical and methodological aspects of the classification of coherence relations and connectives, Levels of Representation in Discourse, *Working Notes of the International Workshop on Text Representation*, Edinburgh University, 3-12. of Dutch and English connectives). Acceptability of the substitution is dependent on the degree of similarity between the two connectives. This leads to various possibilities, from connectives that would be totally overlapping in meaning to partially overlapping, and exclusive ones. An interesting aspect of the substitution test is that it can help uncover the specific meaning of a given connective, e.g. when the substitution of one connective by another in a compatible context leads to systematic (even slight) divergences in meaning of the overall segments, or when substitution is systematically impossible.

While this substitution test can bring us closer to a linguistically motivated classification of connectives because it groups together surface cues that are close in meaning, the main problem is that the whole classification is based on the analysts' judgement. And these judgements may diverge between different analysts, especially when the connectives are largely overlapping in meaning. Furthermore, the substitution test does not provide any help in determining the specific meaning divergences.

The paraphrase test is a frequently used method to analyze the meaning of connectives (Pander Maat & Sanders 1995; Degand 1996, 1998; Pit, Pander Maat & Sanders 1997). Paraphrases help in determining what the meaning of the possible relation is, given the to be connected segments i.e. when the influence of the connective itself is left out. Examples of possible paraphrases are:

Semantic, causal, volitional: *this action is the consequence of the following situation...*

Pragmatic, causal, epistemic: *one may conclude this on the basis of the following situation ...*

The advantage of a paraphrase test is that it forces analysts to be explicit about the interpretation of the meaning of the connective under investigation. Once the connective is left out, the analyst checks whether the segments can be related by the given paraphrase(s) without changing the reading of the overall fragment. Of course, the analyst's intuition is still the ultimate criterion to classify the connective as expressing one or another meaning, but the paraphrase test makes the different interpretations explicit, and thus helps the analyst to systematically check intuitions. Formulation of the paraphrases themselves is based on general theoretical descriptions of the possible meanings and functions connectives can have.

In how much progress does this paraphrase test result in terms of concrete research ? We have an ideal opportunity to answer that question. We can compare the results of two empirical studies, done independently, both making use of this test, and both on the categorization of the Dutch causal connective *dus* ('so'). Table 1 displays the results from these studies (Pander Maat and Sanders, 1995 and Degand, 1996).

DUS	Content-vol	Content-nvol	epistemic	speech-act	textual	Ambiguous
PM & S	8%	0%	58%	0%	2%	16%
Degand	36%	34%	14%	12%	4%	0%

Table 1. A comparison of the relational distribution of *dus* in newspaper corpora

Why are there such enormous differences? A first reason could be that the paraphrase is not formulated in a way that is precise and explicit enough to be applied unambiguously, i.e. what exactly is *conclude*, ..., what exactly is a *volitional* action? There are typical cases in which one may favor a non-volitional content or rather an epistemic interpretation. Very clearly, Pander Maat and Sanders

Degand, Liesbeth & Sanders, Ted (1999) Causal connectives in language use. Theoretical and methodological aspects of the classification of coherence relations and connectives, Levels of Representation in Discourse, *Working Notes of the International Workshop on Text Representation*, Edinburgh University, 3-12. (1995) have given priority to an epistemic categorization, while Degand (1996) has favored classification into the non-volitional content domain.

Second, and perhaps more interestingly, Degand had an additional criterion for the category of epistemic relations: the direction of the reasoning pattern, i.e. it had to concern abductive reasoning in which one reasons from result and (causal) rule to circumstance, without mentioning the rule explicitly, as in *The river is high, so it has been raining a lot*. In fact, distinguishing between different types of reasoning patterns explains why Degand has a smaller category of epistemic causals, as well as a more domain-specific categorization of Dutch causal connectives. As a matter of fact, in this view, only a very restricted number of connectives can express epistemic (abductive) relations.

At this point we want to use the insights from this comparison of the paraphrase studies to get a better hold on a dominant question: What does this tell us about the characteristics of clear domain-specific connectives? Rather than concluding that epistemic relations are restricted to abductive reasoning, we would consider them to be a prototypical type. Every relation would have a prototypical linguistic realization. The question is what this prototypical realization would consist of? A (prototypical) connective is of course part of it, but other elements also seem to co-occur systematically when expressing a given type of relation: perspectivizing elements. Our previous experience has taught us that the substitution and paraphrase tests are not satisfying in getting a hold on these contextual elements. In our view, what is needed here are systematic distribution analyses in corpora. This is exemplified in the following section.

4. Perspectives and prototypical meaning

In the previous section, we came to the conclusion that (the majority of) connectives are not domain-specific, but rather that they could have a *prototypical* (relational) meaning. Our claim here is that this prototypicality is mirrored in the co-occurrence pattern of the connectives, i.e. connectives appear in a typical environment. This environment can be characterized in terms of perspectives.

The core idea of perspective, we want to build on here is the question of who is responsible for the content of the related segments, i.e. from whose perspective is the information conveyed? Two aspects play a primary role here: (i) Is the perspective that of the author or that of another participant? (ii) Is this participant explicitly mentioned in the segment, or is it left implicit? Previous studies have already indicated that connectives tend to show a preference for certain types of perspectivization patterns. In this paper, we would like to refine this question in terms of prototypicality and investigate whether prototypical and non-prototypical uses of connectives are related to different perspectives. In particular, we will concentrate on the Dutch causal connectives *want* and *omdat* ('because'), which have been shown to express different prototypical meanings (Degand, 1996, 1998; Pit et al., 1997). This is also confirmed by our own analysis of 50 random occurrences of the two connectives in a Dutch newspaper corpus. Table 2 shows that there is a clear tendency for *omdat* to express content relations and for *want* to express epistemic relations ($\chi^2_{(2)}$: 30.922; $p < .0001$). This goes on a par with the modality of the first segment of the relation ($\chi^2_{(3)}$: 32.647; $p < .0001$), opinions, and facts or actions in S1 being indicative for epistemic and content relations, respectively. This is shown in Table 3.

Relation	Connective	
	<i>Omdat</i>	<i>Want</i>
Non-vol. causal	12	1
Vol. causal	27	11

Degand, Liesbeth & Sanders, Ted (1999) Causal connectives in language use. Theoretical and methodological aspects of the classification of coherence relations and connectives, Levels of Representation in Discourse, *Working Notes of the International Workshop on Text Representation*, Edinburgh University, 3-12.

Causal epistemic	11	38
	50	50

Table 2: Relational distribution of the connectives *want* and *omdat*.

Degand, Liesbeth & Sanders, Ted (1999) Causal connectives in language use. Theoretical and methodological aspects of the classification of coherence relations and connectives, Levels of Representation in Discourse, *Working Notes of the International Workshop on Text Representation*, Edinburgh University, 3-12.

Modality	Connective	
	<i>Omdat</i>	<i>Want</i>
Opinion	12	38
Action	27	12
Fact	11	0
	50	50

Table 3: Modality of segment 1 in *want* and *omdat* relations.

Furthermore, Table 3 confirms that the two connectives are globally characterized by different perspectivization patterns. To build the perspectivization pattern, we determined first the ‘primary participant’ for each of the related segments, that is, we determined the participant who carries primary responsibility for the content of the segment either as “concluder”, “experiencer”, “observer” or “actant” (see also Degand & Pander Maat, this volume). In addition, we stipulated whether this primary participant was explicitly mentioned in the segment or left explicit, and whether other explicit or implicit perspective indicators¹ were present. At least for segment 1, the connectives are significantly differentiated for all three of these aspects: identity of the responsible participant ($\chi^2_{(5)}$: 10.010; $p < .01$), explicit or implicit realization of this participant ($\chi^2_{(3)}$: 34.114; $p < .0001$), and presence of perspective indicators ($\chi^2_{(2)}$: 41.384; $p < .0001$). Contrary to segment 1, the second segments in *omdat* and *want* relations were significantly different for none of these parameters.

		Omdat n=50	Want n=50
Participant S1	Author	8	25
	Cited sp.	29	14
	Impl. Cit. sp.	2	1
	Generic	2	0
	3 rd pers. pron.	8	3
	3 rd pers. nom.	11	7
Realization participant S1	explicit	32	16
	Impl. Agent	2	0
	Impl. Concl.	7	34
	Absent (fact)	9	0
Perspective indicator S1	Explicit	11	10
	Implicit	7	36
	No P.I.	32	4

Table 4: Perspectivization pattern of segment 1 in *omdat* and *want* relations

Now, what happens if we refine these results taking apart the prototypical (i.e., *omdat* in a content

¹ Other explicit perspective indicators include ‘aanhalingstekens’, verbs of saying or mental processes; implicit perspective indicators include scalar and evaluating adjectives, modality verbs, etc.

Degand, Liesbeth & Sanders, Ted (1999) Causal connectives in language use. Theoretical and methodological aspects of the classification of coherence relations and connectives, Levels of Representation in Discourse, *Working Notes of the International Workshop on Text Representation*, Edinburgh University, 3-12.

relation, *want* in an epistemic relation) and non-prototypical (i.e., *want* in a content relation, *omdat* in an epistemic relation) uses of the two connectives? Our basic hypothesis is indeed that perspectives come on a par with the prototypical meanings and that this should be reflected by a specific perspectivization pattern. This leads to the following hypotheses for the prototypical and non-prototypical uses of the connectives *want* and *omdat*.

- I. In its prototypical use, *want* should favor an (implicit) author perspective.
- II. In its prototypical use, *omdat* should favor an (explicit) 'other participant' perspective or a lack of perspective ('neutral' perspective or fact).

The first hypothesis is justified by the fact that epistemic relations play in the mental space of the participant. Clearly, it is easier to have access to its own mental space (that of the author) than to that of another participant. In addition, since the default is to convey opinions from ones own rather than from other people, perspective indicators may be left implicit in these cases. Implicit author perspectives induce a subjective reading to the segments, i.e. the distance between the participant (the author) and the information conveyed is small. This is also one of the effects of using *want* (see below). As for the second hypothesis, content relations are more objective. They create a distance between the 'responsible' participant and the content of the segment. This should be reflected, either by the lack of perspective ('neutral' perspective or facts) in *omdat*-segments, or by the explicit mentioning of the responsible participant as being different from the author.

On the basis of this, we can now formulate our hypotheses for the non-prototypical cases:

- III. If *omdat* does not express a content relation, we expect there to be (explicit) perspective indicators at another level, i.e. indicating 'epistemicity'.
- IV. If *want* does not express an epistemic relation, we expect there to be (explicit) 'neutralizing' indicators at another level, i.e. distance creating features indicating that the segment is not an opinion or stressing its factual aspects.

In regard to these hypotheses, we expect the data in Table 3 above, which are not in agreement with the prototypical hypothesis, to be on the account of non-prototypical uses. To test this, we segmented the 50 *omdat* and *want* occurrences into prototypical and non-prototypical cases. The results are recapitulated in Table 4.

		Omdat		Want	
		Prototyp. n=39	Non-prototyp. n=11	Prototyp. n=38	Non-prototyp. n=12
Modality S1	Opinion	1	11	38	0
	Action	27	0	0	12
	Fact	11	0	0	0
Participant S1	Author	7	1	25	0
	Cited sp.	11	7	12	2
	Impl. Cit. sp.	0	2	1	0
	Generic	2	1	0	0
	3 rd pers. pron.	8	0	0	3
	3 rd pers. nom.	11	0	0	7

Degand, Liesbeth & Sanders, Ted (1999) Causal connectives in language use. Theoretical and methodological aspects of the classification of coherence relations and connectives, Levels of Representation in Discourse, *Working Notes of the International Workshop on Text Representation*, Edinburgh University, 3-12.

Realization participant S1	explicit	26	6	4	12
	Impl. Agent	2	0	0	0
	Impl. Concl.	2	5	34	0
	Absent	9	0	0	0
Perspective indicator S1	Explicit	4	7	7	3
	Implicit	3	4	31	5
	Absent	32	0	0	4

Table 4: Perspectivization pattern of segment 1 in prototypical and non-prototypical uses of *omdat* and *want*

First of all, it should be noted that the connectives *omdat* and *want* differ significantly from one another both in their prototypical and their non-prototypical use, and this on all aspects of the pattern. This simply confirms that the two connectives occur in a systematically different environment². Actually, the table shows that hypotheses I and II concerning the prototypical use of *omdat* and *want* are borne out. That is, prototypical *omdat* shows a clear preference for a third person participant (82%), the participant is explicit in 66% of the cases or absent (23%), and in 82% of the cases there are no other perspective indicators. In contrast, prototypical *want* has a majority of author perspectives (65%) or cited speakers (31%); this participant is left implicit in 89% of the case; and most of the time there are other implicit perspective indicators (81%). So, we can now affirm that examples (7) and (8) are prototypical for *omdat* and *want*, respectively.

- (7) In Istanbul waar het sinds vrijdag sneeuwt, ligt een vijftig centimeter dikke laag. Het vliegveld moest vanmiddag twee uur worden gesloten *omdat* de landingsbanen met ijs bedekt waren.
 ‘In Istanbul, where it has been snowing since Friday, there’s a snow layer of 50 centimeters. The airport had to be closed for 2 hours this afternoon because the landing runways were covered with snow.’
- (8) In zijn vrije tijd bestudeerde hij de werken van Haydn en Mozart. Dat was voor die tijd nogal uitzonderlijk, *want* deze twee componisten waren toen in Italië nauwelijks bekend.
 ‘During his spare time he studied the works of Haydn and Mozart. That was quite exceptional at that time, because the two composers were then hardly known in Italy.’

The interesting question, however, is whether the prototypical and non-prototypical uses of one connective are also significantly different from one another, because only then will it be possible to state that the prototypical meaning of a connective is linguistically reflected in its immediate environment. And indeed, this is the case. For all aspects of the perspectivization pattern of segment 1, the prototypical and the non-prototypical uses are significantly different both for *omdat* and *want*. This means that a connective tends to favor a specific environment in its typical use. In terms of methodology, this pleads against the substitution test, where the proper meaning of the segments is not taken into account.

If we now have a more detailed look at the non-prototypical cases, we can observe that hypothesis III concerning the non-prototypical use of *omdat* is verified too. There are indeed a number of explicit

² There is however one exception, the presence of perspective indicators in the non-prototypical use of the two connectives is 'almost' significant ($\chi^2_{(2)}: 5.678; p = .058$), while all other features are clearly significant.

Degand, Liesbeth & Sanders, Ted (1999) Causal connectives in language use. Theoretical and methodological aspects of the classification of coherence relations and connectives, Levels of Representation in Discourse, *Working Notes of the International Workshop on Text Representation*, Edinburgh University, 3-12.

perspective markers indicating that the relation has to be interpreted at 'another' level, namely the epistemic level. This appears from the fact that the modality of segment 1 is always an opinion. There is a strong tendency for the 'responsible participant' to be a cited, i.e. clearly identifiable, speaker (72%) that is either explicitly mentioned in the segment (54%) or signaled by means of explicit perspective indicators (63%). In other words, the connective *omdat* can be pulled in the direction of *want*, provided this is explicitly signaled by some linguistic means in the segment. At the same time, this explicit markers have an objectifying or factualizing effect, i.e. the use of *omdat* together with these explicit perspective indicators keeps the author at some distance from the information conveyed. A typical example is given in (10).

(10) Volgens prof. E. te Velde zal de vraag naar kunstmatige bevruchting en vruchtbaarheidsonderzoeken stijgen *omdat* steeds meer vrouwen op latere leeftijd zwanger willen worden.

‘According to prof. E. te Velde the demand for artificial insemination and fertility analyses will be rising because more and more women want to get pregnant at a later age.’

For the non-prototypical use of *want*, the hypothesis is only partially confirmed. The modality of segment 1 is always an action (content relation) and this is reflected by the fact that the responsible participant is nearly always a third person (nominal or pronominal) (83%), i.e. the agent of the action, who is explicitly mentioned in the segment (100%). However, there are no clear other perspective indicators that would favor the interpretation of the segment at a content level. Actually, from an intuitive point of view, these *want* cases still carry some 'epistemic flavor'. This is also reinforced by the fact that in these cases, segment 2 expresses most often an opinion (66%) marked by implicit perspective indicators (66%), with a tendency for the responsible participant to be left implicit (50%). This is recapitulated in Table 5. Actually, in these atypical *want* cases, it is as if the opinion of the actant introduced in segment 1 is implicitly given in segment 2. The use of *want* together with the implicit perspective give a subjective reading to the segment. It is as if the author gives the reader a look in the head of this other participant. This provokes a small distance, or even identification between author and participant. In other words, it is as if *want*, when it is used to express a content relation, imports some of its epistemic meaning. An example is given in (11), where the second segment should clearly be interpreted as an opinion of the participants in the first segment, and not of the author.

Perspective		Want
Modality S2	Opinion	8
	Action	0
	Fact	4
Real. PP S2	Explicit	2
	Implicit	6
	Absent	4
P.I. S2	Explicit	3
	Implicit	8
	No P.I.	1

Table 5: Perspective of segment 2 in non-prototypical *want* relations

(11) Over voorkeuren voor burgemeesters laten bodes zich overigens niet uit. *Want* ook diplomatie hoort bij het vak, net als "enige afstand tot de baas".

‘The servants don’t say anything about the preferences of the burgomaster. Because diplomacy is

Degand, Liesbeth & Sanders, Ted (1999) Causal connectives in language use. Theoretical and methodological aspects of the classification of coherence relations and connectives, Levels of Representation in Discourse, *Working Notes of the International Workshop on Text Representation*, Edinburgh University, 3-12.
part of the job, as well as “some distance to the boss”.’

Concluding remark

What we may conclude from this is that connectives are not domain-specific in their use, but that they have a prototypical meaning. Actually, this prototypical meaning overlaps with the classical relational domains or categories. This explains why most scholars agree on the categorization of these connectives.

Our analysis has furthermore shown that this prototypical meaning comes on a par with a specific co-occurrence pattern, i.e. when a connective is used prototypically, the segments it relates (at least the first segment in backward causal connectives) also show some recurring pattern. We defined this pattern in term of perspectives. When a connective is used in an atypical way, this perspective pattern changes. The change of perspective (and perspective indicators) is a signal to the hearer/reader that the connective should not be interpreted in its “classical” way. Nevertheless, even when used atypically, connectives keep something of their kernel meaning that they “import” into their new discourse environment. These findings are compatible with Degand and Pander Maat’s (this volume) proposal for a scalar classification of connectives.

Acknowledgements

Liesbeth Degand is a researcher at the Belgian National Fund for Scientific Research (F.N.R.S). She is also supported by a “Pôle d’attraction interuniversitaire” (PAI-4/19) financed by the Ministry of Scientific affairs of the Belgian federal government.

References

- Bateman J.A. & Rondhuis, K.J. (1997). “Coherence relations’: towards a general specification”. *Discourse Processes* 24: 3-49.
- Degand, L. (1996) *A situation-based approach to causation in Dutch with some implications for text generation*, Doctoral Dissertation, Université catholique de Louvain, Louvain-la-Neuve.
- Degand, L. (1998) Het ideationele gebruik van *want* en *omdat*: Een geval van vrije variatie? *Nederlandse Taalkunde* 4, 309-326.
- Degand, L. & Pander Maat, H. (this volume) Scaling causal relations in terms of Speaker Involvement.
- Hovy, E. (1990) “Parsimonious and profligate approaches to the question of discourse structure relations”. In *Proceedings of the 5th International Workshop on Natural Language Generation*, 128-136.
- Knott, A., & Dale, R. (1994). Using linguistic phenomena to motivate a set of coherence relations. *Discourse Processes*, 18, p. 35-62.
- Knott, A. (1996). *A data-driven methodology for motivating a set of coherence relations*. Doctoral dissertation, University of Edinburgh, Edinburgh, Scotland.
- Knott, A. & Sanders, T. (1997). The categorization of coherence relations and their linguistic markers: an exploration of two languages. *Journal of Pragmatics*, 30, 135-175.
- Martin, J.R. (1992). *English text: systems and structure*. Amsterdam: Benjamins.
- Moore, J.D. & Pollack, M.E. (1992). A problem for RST: The need for multi-level discourse analysis. *Computational Linguistics* 18: 537-544.
- Pander Maat, H. (1998). The classification of negative coherence relations and connectives. *Journal of Pragmatics*, 30, 177-204.

- Degand, Liesbeth & Sanders, Ted (1999) Causal connectives in language use. Theoretical and methodological aspects of the classification of coherence relations and connectives, Levels of Representation in Discourse, *Working Notes of the International Workshop on Text Representation*, Edinburgh University, 3-12.
- Pander Maat, H. & T. Sanders (1995). Nederlandse causale connectieven en het onderscheid tussen inhoudelijke en epistemische coherentie-relaties. *Leuvense Bijdragen*, 349-374.
- Pit, M. Pander Maat, H. & Sanders, T. (1997). *Doordat, omdat en want*. Perspectieven op hun gebruik., *Taalbeheersing* 19: 238-251.
- Sanders, T. (1997). Semantic and pragmatic sources of coherence; On the classification of coherence relations in context. *Discourse Processes*, 24, 119-147.
- Sanders, T. & W. Spooren (1999). Communicative intentions and coherence relations. To appear in: W. Bublitz, U. Lenk & E. Ventola (red.) *Coherence in text and discourse*. Amsterdam / Philadelphia: Benjamins.
- Sweetser, E.E. (1990). *From etymology to pragmatics*. Cambridge etc.: Cambridge University Press.