Ted Sanders and Wilbert Spooren

Exceptions to rules: a qualitative analysis of backward causal connectives in Dutch naturalistic discourse

Abstract: Language users systematically prefer one lexical item (*because*) over another (even highly similar) one (*since*) to express a causal relationship in discourse. Such choices provide a window on speakers' cognitive categorizations, and have been modeled in previous work in terms of subjectivity. This paper analyzes the Dutch connectives *omdat* ('because') and *want* ('since/for') in written text, conversation, and chat interactions. These can be considered a case in point for linguistic categorization since related European languages show similar distinctions. We sketch a profile for the interpretation of *omdat* and *want* based on corpus analyses of large numbers of occurrences in different media and genres. However, we focus on the deviations from the prototypical use of the connectives. We analyze instances of those deviations, in order to be able to understand the deviating use. We conclude that deviations should be interpreted in terms of core elements of the prototypical use. Therefore, the semantic-pragmatic profile of *want* and *omdat* should not be considered as hard-wired all-or-nothing rules, but rather as a prototype structure with a core meaning/use and more peripheral uses. The non-prototypical, peripheral uses are motivated deviations: we need the elements in the core profile to understand the deviations.

Keywords: causality; connectives; subjectivity; Dutch; discourse; prototype structure; coherence.

━━━
Ted Sanders: Utrecht Institute of Linguistics, Universiteit Utrecht, Trans 10, Nl- 3512 JK Utrecht, The Netherlands. E-mail: T.J.M.Sanders@uu.nl
Wilbert Spooren: Centre for Language Studies, Radboud University, Nijmegen, The Netherlands. E-mail: W.Spooren@let.ru.nl

1 Introduction

In everyday communication, language users regularly express some sort of causal relation. While telling a story, they may want to express the reason why an event took place, why something happened, or why a certain situation is the case, as in

(1). When expressing their opinion, they may provide evidence in favor of their point of view, as in (2).

(1) The neighbors are not at home because they are both at the office.

(2) The neighbors are not at home, because the lights are off.

In English, both relations can be made explicit with the connective *because*. This shows how these relations, irrespective of their obvious differences, share the conceptual relation of causality. As noted by many linguists over the years, the relations clearly differ, too: in (1) the causality is located on the level of the state-of-affairs in the world: an explanation why the neighbors are not there. In (2), the *because*-clause does not present the reason why the neighbors are out, but rather the reason for the speaker to utter this claim: "I am sure they are not at home, because…". This systematic distinction between types of causal relations has been described in terms of dichotomies like semantic–pragmatic, internal–external, locutionary–illocutionary, subject matter–presentational, or objective–subjective.

 Speakers of English can use the connective *because* to express both (1) and (2), even though they could also use *since* or *for* in example (2). However, in various other European languages, there is not one, general causal connective, but rather two more or less "specialized" connectives. In Dutch, for instance, the connective *want* ('because/for') would be used to express the causality in (2), which could simply not be expressed using *omdat* ('because') (De Vries 1971; Verhagen 2005; Sanders and Spooren 2009). *Omdat*, on the other hand, would be used to express example (1). Prototypical *want*-examples are (3), taken from a corpus of spoken Dutch, and (4) from a newspaper corpus.

(3) [*S1 dat is gewoon krankzinnig*].
     [S1 that is simply insane].
     want [*S2 als hij uhm mensen goed inschat moet ie ook weten dat ik m'n uiterste best doen om dat zo snel mogelijk voor elkaar te krijgen*].
     because [S2 if he uhm is such a good judge of character then he should also know that I am doing my very best to take care of that as soon as possible].

(4) [*S1 Drie vrouwen van middelbare leeftijd worden achterna gezeten*] *omdat* [*S2 ze het waagden te protesteren*]
     [S Three middle-aged women are chased] because [S2 they dared to protest]

Highly similar observations are available for German *weil* versus *denn*, and French *parce que* versus *puisque* and *car*. By systematically choosing one item over another, people distinguish between several types of causality. Hence, such

choices provide a window on speakers' cognitive categorizations of causality. The linguistic study of the meaning and use of causal connectives may reveal insights into human categorization of causality (Sanders and Sweetser 2009).

Since the mid-1990s we have witnessed a rise in corpus studies to investigate the organization of the lexicon of connectives in several languages, seeking to find the system behind the meaning and use of (causal) connectives, which has led to empirical tests in actual language use (see Sanders and Spooren [2009] and other contributions to Sanders and Sweetser [2009]).

Most of the corpus studies have used data from written corpora. There is a certain urgency to add other data to the empirical foundation of theories on the categorization of connectives. Several studies of spontaneous conversations suggest a typical usage pattern of causal connectives in conversations. Günthner (1993) and Keller (1995) demonstrated that German *weil* can express epistemic relations (like [2]) in spontaneous conversations, whereas in written language it seems to be reserved for what Sweetser (1990) has labeled the content domain, as in example (1).[1] Similarly, Zufferey (2010) concludes that French *puisque* has a strong preference for epistemic use in telephone conversations. Such results show that written language as the basis for analysis may lead to a distorted picture. A principled point is that written language deviates from the prototypical communicative situation that spontaneous conversations provide in several respects (Clark 1996).

We focus on the meaning and use of *want* and *omdat* as a case in point of how European languages encode backward causal relations that differ in subjectivity. In most corpus studies, a difference has been found between connectives used in an objective (content) context and those used in subjective (epistemic) contexts, but this use is generally characterized in terms of tendencies rather than in a black-and-white fashion. Many studies of French, German, and Dutch causal connectives have drawn attention to the fact that typically objective connectives sometimes also occur in subjective contexts (and vice versa) (Stukker et al. 2009 and the references cited there). How should these apparent "counterexamples" against the categorization hypothesis be accounted for? In this paper, we explore the linguistic categorization of causal relations by analyzing apparent counterexamples in terms of subjectivity. We show that objective causality and subjective

---

1 According to several German linguists, a diachronic change has recently taken place in the meaning and use of *weil* and *denn* (Günthner 1993; Keller 1995). Interestingly, *denn* seems to be largely replaced by a subjectively used *weil*. Originally, *weil* is a subordinate connective expressing objective causality. The new *weil*, however, which seems to be taking the position of the coordinate connective *denn*, also seems to take *denn*'s main clause word order (Keller 1995). This development was found in the analysis of spontaneous discourse (Günthner 1993; but see Wegener 2000).

causality function as the connectives' prototypical contexts of use, whereas the apparent "counterexamples" can be interpreted as less prototypical instances (Stukker and Sanders 2012).

2 The characterization of *want* and *omdat* in terms of subjectivity

Subjectivity is a prominent notion in functional and cognitive linguistics. We adopt an integrative approach to subjectivity. In line with earlier work on causal connectives (Pander Maat and Sanders 2000), we define an utterance as subjective when its interpretation requires an active *Subject of Consciousness* (from now on SoC). That is, we consider an utterance subjective if it is tied up to a subject experiencing feelings and emotions, being an agent, taking epistemic stances, etcetera. Consequently, a SoC crucially involves an animate subject, a person, whose intentionality is conceptualized as the ultimate source of reasoning, evaluating, describing, or acting "in the real-world." In other words, an utterance is subjective because there is some thinking entity in the discourse who evaluates. For instance, "He thought the bird was beautiful" is subjective because it involves an evaluation by a character in the discourse. Compare this with an utterance like "A kingfisher is a bird," which is presented as a fact in the world that does not depend on the evaluation by an SoC. To be more precise, in the utterance *I thought the bird was beautiful* the validity of the proposition "The bird is beautiful" depends on the SoC *I*, whereas in the utterance *A kingfisher is a bird* the proposition "A kingfisher is a bird" can be verified directly in the non-linguistic reality.

Authors/speakers can be SoCs, but characters can also function as such. The author/speaker is the first voice in the discourse, who has constant access to her feelings and thoughts. She does not have access to the feelings and thoughts of a third person. As a result, *I think the bird is beautiful* can be a direct report of an inner feeling, whereas *he thinks the bird is beautiful* is a description of an evaluation. This implies that first-person evaluations are more subjective than third-person evaluations. The difference, then, between the speaker/writer versus a character as SoC is that the first type concerns a first voice, which is grounded in the deictic center of communication (Sanders et al. 2009; Sanders and Spooren submitted). This resembles Traugott's (1989, 1995) view on subjectivity as closeness to the communicative "here and now": the speaker here and now asserts that a particular state of affairs holds. By contrast, character subjectivity concerns a third person in the discourse, which is more distant from the deictic center of communication.

The examples mentioned so far, be it first-person SoCs or third-person SoCs, are explicit descriptions of evaluations and consequently they are more or less objective. In terms of Sweetser (1990) this type of subjectivity may still be in the content domain. Yet, evaluations are often much more implicit. Especially when the speaker/author is first-person SoC and the evaluation concerns the here and now, spontaneous evaluations typically are of the type *Paris is great*, i.e., a first-person SoC, the utterance expresses an evaluation and the SoC remains implicit. The most subjective types of utterances are those in which the speaker is SoC in first person, but remains implicit, off-stage (Langacker 1990).

In our analysis so far we have presented subjectivity as a property of utterances. However, subjectivity can also reside in the nature of relations between utterances: the link can be of a content, epistemic, or speech-act type. In fact, it is this kind of subjectivity that is the main focus of our interest, since we are interested in subjectivity between utterances on the discourse level (Sanders et al. 1992). Causal links in the content domain (example [1]) are objective. Epistemic causality is inherently subjective because the speaker is actively reasoning toward a conclusion or concluding something on the basis of an observation (example [2]) in the here-and-now. Cases of speech act causality (example [7]) are also subjective: the speaker is performing a speech act and motivating that act on the basis of an observation.

In our analyses of corpus examples, we have taken the relational nature of subjectivity into account in two ways: we distinguish between different types of causal relation, and we specify the SoC responsible for the causal link. See the following examples, from Sweetser (1990).

(5) John came back because he loved her.

(6) John loved her, because he came back.

(7) What are you doing tonight, because there's a good movie on.

Example (5) is a content relation, with a third-person SoC, *John*. Examples (6) and (7) are of an epistemic and speech-act nature, respectively, with the speaker as the SoC. Note that this does not mean that there is a perfect correlation between relation type and SoC. Consider example (8), based on the epistemic relation in example (2).

(8) That Saturday evening, they had a huge argument. His wife was saying the neighbors had returned from their vacation, but Daan disagreed. [S1 Surely the neighbors were not at home], because [S2 the lights were off.]

Example (8) is an epistemic relation: the SoC concludes on the basis of an argument that the neighbors are not at home. What is special about this example is that the SoC is not the speaker (as in examples [2] and [6]), but a third person (Daan). This is a case of free indirect speech, and shows that epistemic relations can occur in a third-person SoC context.

In sum, we have decomposed the complex construct of subjectivity in terms of four characteristics of causal connections. In a recent study (Sanders and Spooren submitted), we have investigated the extent to which these characteristics co-occur. We analyzed a corpus of *omdat-* and *want*-cases from written, spoken, and chat discourse. Our main hypothesis was that *want* occurs in more subjective contexts than *omdat*, irrespective of the medium. We formulated four specific hypotheses on the way in which the connectives *want* and *omdat* would show differences in terms of subjectivity. The main results of our corpus research are summarized in Table 1.

Table 1: Summary from Sanders and Spooren (submitted) of quantative corpus analyses in three different media (spoken, written, chat) (percentages)

	Relation type		Prop. attitude		Type of SoC		Marking of SoC	
	content	Ep/SA	evaluation	other	1/2 per	3rd per	explicit	implicit
omdat	91.6	8.4	38.8	61.2	61.0	39.0	62.8	37.2
want	38.1	61.9	53.6	46.4	80.9	19.1	49.2	50.8

Note: Ep = epistemic; SA = speech act; Prop. attitude = propositional attitude; 1/2 per = first and second person; 3rd per = third person

These results show that the difference between *omdat* and *want* is not only a difference in the type of coherence relation but also in a difference in the other characteristics of subjectivity: Across all media *want* is used more often
–   to express subjective relations (epistemic, speech act) than *omdat*;
–   to support a judgment than *omdat*;
–   with first- and second-person SoCs than *omdat*;
–   with an implicit SoC than *omdat*.

These results allow us to sketch the following prototypical semantic–pragmatic profiles for the two connectives.

Want instructs the reader to find the nearest SoC, and to create a causal connection "P → Q" between S2 (expressing P) and S1 (expressing Q); the SoC is respon-

sible for this connection. The connection is a non-content relation: epistemic (Claim–Argument, Evidence) or speech act.

Omdat instructs the reader to create a causal connection "P → Q" between S2 (expressing P) and S1 (expressing Q). The connection is a content relation (volitional or non-volitional cause).

In this paper we will analyze discourse fragments from our corpora that display uses that deviate from these profiles. Our aim is to identify regularities behind those deviations.

3 Data and methodology

We used data from three different types of corpora, in different media: a written corpus, a spoken corpus, and a chat corpus. These corpora vary with respect to the degree of production planning (carefully edited in the written corpus, spontaneous in the spoken and chat corpus), immediacy of response (delayed in the written corpus, intermediate in the chat corpus, and direct in the spoken corpus), and degree of integration (highly integrated in the written corpus, fragmented in the chat corpus and the spoken corpus) (see Chafe [1994] and Clark [1996] for a discussion of such dimensions of variation between various modes of communication). Together they provide an interesting spectrum of variation in the contexts in which *want* and *omdat* can occur.

For the written medium we made use of the pilot version of the D-COI corpus (D-COI 2006). We randomly selected 100 occurrences of *omdat* and 100 occurrences of *want*. For the spoken medium we made use of the Corpus of Spoken Dutch (Corpus Gesproken Nederlands, CGN; Oostdijk 2000). From the spontaneous conversations and interviews in this corpus we randomly selected 100 fragments with *want* and 100 fragments with *omdat*. For the chat medium we have used the VU Chat corpus (Spooren et al. 2010) and the pilot version of the CONDIV corpus (Grondelaers et al. 2000). Because these chat corpora are fairly limited in size and *omdat* is not used very frequently in chat, we analyzed 51 occurrences of *omdat* and 100 occurrences of *want*.

For each occurrence of *want* and *omdat* an analysis was made in terms of the profile described in the previous section. That is, for each occurrence it was established what was the propositional attitude of the first segment; whether there is a Subject of Consciousness (SoC); who it is; whether it is referred to explicitly; and finally, what type of coherence relation is expressed by the connective. To make these decisions, we made use of a codebook that is available upon

request from the authors. We made use of paraphrasing and substitution as strategies for analysis (see Knott and Dale [1994] and Pander Maat and Sanders [2000] for these strategies). The reliability of the analysis was guaranteed, on the one hand by the explicit indications for coding decisions in the codebook, and on the other by following what Spooren and Degand (2010) have named a complete double-coding strategy: the two authors both coded all fragments and discussed discrepancies. Occasionally the authors did not reach agreement, in which case we concluded that more than one analysis of a particular fragment was possible.

All in all this led to a corpus of 251 *omdat*-fragments and 300 *want*-fragments, which were coded for the four parameters discussed earlier. As explained earlier, we expected the *want*-fragments to be more subjective than the *omdat*-fragments. The quantitative analysis of this corpus we have reported elsewhere (Sanders and Spooren submitted). The analyzed corpus allowed us to identify fragments that completely fit the theoretical profiles of *want* and *omdat*. But it also allowed us to identify fragments that do not fit the profile. Here we will focus on these deviant cases.

In the most extreme case a fragment deviated from the profile in all four aspects. We will discuss such non-prototypical cases. We have selected a number of *want*-fragments that in several respects do not seem to be subjective, and *omdat*-fragments that do not seem to be objective. We will discuss these cases in a qualitative way, in order to be able to understand why the fragment contains *want* or *omdat*. For this we will make use of whatever information is available. For example, for the written corpus we sometimes made use of newspaper articles referred to in a text; for the spoken corpus we sometimes made use of available audio fragments.

4 *Omdat* in subjective contexts

In this section we discuss a number of *omdat*-fragments which deviate from the profile sketched earlier in that they seem to be more subjective than commensurate with the profile.

Our first example is (9), taken from a background article on the possibilities of prosecution of the Surinam leader Desi Bouterse (NRC-Handelsblad, 7 March 2000).

(9) *En ook Th. van Boven, hoogleraar internationaal recht aan de Universiteit Maastricht, vindt de beschikking "vanuit het oogpunt van de mensenrechten*

alleen maar toe te juichen". Maar de verdere uitwerking noemt Van Boven nog niet zo eenvoudig: "Niet voor niets spreekt het hof geen definitief oordeel uit omdat een volkenrechtdeskundige nog eens naar een aantal wezenlijke punten moet kijken. "[*S1 Dat is heel verstandig*]*, OMDAT* [*S2 de internationale jurisdictie erg in beweging is*]*"*

And even Th. van Boven, professor of international law at Maastricht University thinks that the decision "can only be welcomed from the point of view of human rights". But Van Boven thinks that the further development will not be that simple: "It is for a good reason that the court does not give a final judgment because a specialist in international law still needs to have a look at some crucial points. "[S1 That is very sensible], OMDAT [S2 the international jurisdiction is very much on the move]"

The context in which *omdat* occurs seems a perfectly suitable environment to use *want*: the first segment expresses a judgment; the SoC is implicit; the SoC is the speaker; and the relation is epistemic (Claim–Argument). The *omdat* relation can be paraphrased as "My claim is that it is very sensible, and my argument is that the international jurisdiction is very much on the move". A content paraphrase (of the type "the situation is that it is very sensible and the cause for that situation is that the international jurisdiction is very much on the move") is impossible in this context. Hence, the example fits the prototypical profile of *want* perfectly. So why did the author use *omdat* instead?

The explanation might be that this is a case of rhetorical use of *omdat*. The author deliberately chooses to express the causality with *omdat* rather than with *want*, to suggest that the conclusion does not follow from a self-constructed subjective line of reasoning, but that it follows more or less deterministically from the circumstances in the situation. This objectifying type of use has been described before in legal environments where judges use objective markers to suggest that it is not they, the judges, who draw the conclusion that the defendant has to be convicted, but that it follows necessarily from the law (Van den Hoven 1997). The processing equivalent of this rhetorical use has been demonstrated in research that shows that readers are sensitive to subjective versus objective marking of coherence relations in persuasive texts: Kamalski et al. (2008) have shown that readers read slower and are less convinced upon reading a persuasive text with subjective markers. The authors explain this effect in terms of forewarning: Readers are alerted that the author is too obviously trying to convince them, and hence resist persuasion.

A use of subjective *omdat* that we encountered more than once in our spoken corpus is exemplified by (10).

(10) [Context: B is a secondary school teacher of Dutch who is interviewed by A
    about his profession)

1   A: *uh wat vindt u van een uh initiatief als Het Groot Dictee Der Nederlandse*
        Taal dat dan uh televisie uitgezonden wordt en zo?
        uh what do you think of an uh initiative like the Great National Dutch
        Dictation that then uh is transmitted [on] television and such?

2   B: xxx

3   B: *een sport*
        a sport

4   A: *een sport?*
        a sport?

5   B: *een sport*
        a sport

6   B: *uhm ik weet niet of dat in het belang is van uh de spelling omdat daar uh*
        ja toch zeer uitzonderlijke woorden worden genomen vaak uh om de
        mensen uh allee te belonen
        uh I do not know whether that is in the interest of uh the spelling [capa-
        bilities] OMDAT they uh yes take very exceptional words often uh to uh
        well reward the people

In (10) it is possible to replace *omdat* by *want*. And that resonates with the rather
subjective environment in which *omdat* occurs. The SoC is the speaker, who is
involved in a reasoning process. Why did the speaker use *omdat*? A possible ex-
planation is that the speaker makes his reasoning process explicit; he objectifies
it, so to speak, also in the Langackerian sense of the word: the speaker puts him-
self on stage by the explicit reference to his thought process *ik weet niet of* ('I do
not know whether'). The act of reasoning, then, resembles volitional actions,
which are typically expressed by *omdat*.

    In our next example, (11), segment 1 expresses a conclusion and segment 2 an
argument. This constitutes a typical subjective epistemic relation, which is typi-
cally expressed with *want*.

(11) *De deelgemeentelijke ambtenaar die ik hierover aanschreef vond ook dat het*
    eigenlijk wel meeviel met de rotzooi. Nu is de situatie zo, dat ik zelf de straat
    veeg. Ik ben genoodzaakt dat vuil zelf in een zak te scheppen. Want de berg die
    wekelijks ontstaat is te groot voor de veegploeg om mee te nemen. [S1 Rotter-
    dam is dus zo vies] OMDAT [S2 de reinigers er geen boodschap aan hebben].

    The borough civil servant to whom I wrote about this matter also was of the
    opinion that the mess wasn't as bad as that. Now, the situation is that I

sweep the street myself. I am forced to put that dirt into a bag myself. Be-
cause the pile that gathers every week is too large for the street cleaners to
take it along [S1 So Rotterdam is that dirty] OMDAT [S2 the cleaners won't
have anything to do with it].

This is an example from a letter to the editor in which a Rotterdam inhabitant
complains about the waste in the neighborhood streets. The substitution with
want is possible.

(11′) Rotterdam is dirty WANT the cleaners won't have anything to do with it.

However, the paraphrase in (11′) neglects the presence of the summarizing *dus*
('so') and the intensifier *zo* ('that'). What is different in (11′) is that it suggests that
the new part of the utterance is "Rotterdam is dirty". In fact, this part of the infor-
mation is already given in the previous context. The real new part of the utterance
is that the author has discovered what the reason is for Rotterdam being dirty: "I
conclude that the reason for Rotterdam's dirtiness is that it is no concern of the
cleaners (and not some other reason)". Therefore, an epistemic conclusion is
drawn here (which is signaled by *dus* 'so'). The content of that conclusion is not
only the first segment S1, but the entire reason relation S1–Connective–S2. A dis-
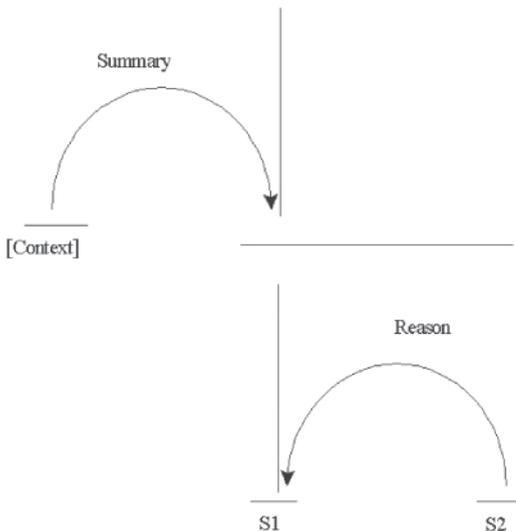course structure of the fragment is given in Figure 1.

Fig. 1: Structure diagram of fragment (11)

This structure is perfectly expressed with *omdat*, whereas *want* (as used in [11']) would have suggested the structure shown in Figure 2.
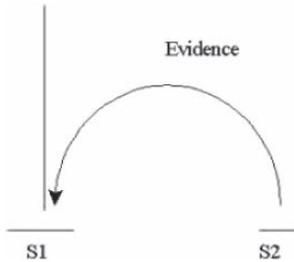

Fig. 2: Structure diagram of the relation in fragment (11')

In short, the writer uses *omdat* in (11) to avoid the suggestion that the writer's conclusion is simply S1, whereas it is in fact the complete causal connection S1–Connective–S2.

We have now seen several uses of *omdat*, which can be considered non-prototypical because they expressed a more subjective relation than expected on the basis of the connective's semantic–pragmatic profile.

In the case of example (11) *omdat* was used to avoid misunderstandings, that is, to ascertain that the evaluation expressed in the discourse is not attributed to the SoC, but to another participant. In another case (9), *omdat* was used to objectify the opinion expressed by the SoC: it was not the individual opinion of the SoC who is responsible for the causality, but rather that the causality follows from a general rule. In short, subjective *omdat* seems to be used to create a distance between the SoC and the opinion expressed: thanks to *omdat* the reader infers that the SoC is not responsible for that opinion. Example (10), from the spoken corpus, is of a different nature, as it is used to signal that the SoC is put on stage explicitly.

5 *Want* in objective contexts

This section analyzes *want*-fragments that at first sight fly in the face of the profile that we sketched previously. The first set of examples all concern a use of *want* that seems to signal empathy with a third-person SoC. Consider example (12), from a letter to the editor (Trouw, September 2002), responding to a newspaper column.

(12) *Negen*

In tegenstelling tot wat Peter Sierksma in zijn column beweert (Boeken, 31 augustus), rust er op negen wél zegen. Niet voor niets wil iedere kleine jongen nummer negen zijn op het voetbalveld. Negen is scoren. [*S1 Voetbalclubs als AZ, De Graafschap en NEC dromen van nummer negen*]. *WANT* [*S2 negen is het linkerrijtje*]. *Ook buiten de sport is de negen interessant.* [. . .]

'Nine
Contrary to what Peter Sierksema is saying in his column (Books, August 31), nine is blessed. It is for a good reason that every young boy wants to be number nine on the football field. Nine means to score. [S1 Football teams like AZ, De Graafschap and NEC are dreaming of number nine]. WANT [S2 nine means the left column]. Also outside of sport the number nine is interesting. [. . .]'

Here, the SoCs are the football clubs, who would dream of the number 9 position in the football league (in the Netherlands the positions of the football teams are typically presented in two columns of nine clubs, so being in the left column implies being in the top of the league). Hence the SoCs are third person, which is not prototypical for *want*. Still, we can understand why the author uses *want* here: first he suggests that he knows what it is that the clubs dream of and he gives their motivation. He does this from the perspective of the clubs, which has the effect of empathizing with them. Consider the difference with (13), the *omdat*-counterpart of (12) we constructed.

(13) Football clubs like AZ, De Graafschap and NEC dream of number nine, OMDAT nine is the left column.

This is a clear description of the reason for the dream rather than standing in the shoes of the clubs and co-experiencing their fantasy. This use of *want* illustrates the well-known use of free indirect speech (FIS; Banfield 1982; Sanders 2010): the second segment is not merely presented from the perspective of the speaker but also from that of a third person. This creates the effect of a double voice (polyphony; Bakhtin 1981; Ducrot 1984). In a recent analysis of this effect of *want* (Sanders, Sanders and Sweetser 2012), it is shown that authors and speakers regularly put this "Easy Identification"-effect of *want* to use. This usage is found in literary novels but also in spoken spontaneous discourse.

In our corpus cases, this FIS character is very clearly demonstrated in fragment (14).

(14) *Verkiezingen*
[*S1 Het kabinet wil de verkiezingen voor de Provinciale Staten en de gemeen-
teraden bundelen, viel onlangs te lezen.*] *WANT* [*S2 de opkomst bij de Staten-
verkiezingen viel tegen.*]

Elections
[S1 The cabinet wants to combine the elections for the provincial and the
municipal councils, as could be read recently.] WANT [S2 the turnout at the
provincial council elections was disappointing.]

By using *want*, the author maintains the perspective of the SoC, the cabinet: We
experience the reasoning by the cabinet "we are going to combine the elections
because of the poor turnout". If we replace *want* by *omdat* the result is an author
perspective, in which the SoC is the author, who describes the cabinet's proposal
from his or her perspective. In short, the use of *want* seems to underline that the
cabinet and not the author is the SoC: it is the cabinet that is responsible for this
reasoning.

Example (15) illustrates a similar point.

(15) A: *maar dat kun je sommige mensen niet uitleggen.*
but it is difficult to explain this to some people.
A: *dan krijg je een tirade over je heen.*
then you'll get a tirade against you.
B: *maar ja jij bent bezig met een ander karwei.*
but well you are busy carrying out another chore.
A: *ja.*
yes.
A: *maar* [*S1 hij eist dat je dat nu doet.*]
but [S1 he demands that you do it right now.]
A: *WANT* [*S2 dat moet nu daar hangen.*]
WANT [S2 that should hang there right now.]

In this example, from the Corpus of Spoken Dutch, two male secretaries are com-
plaining about the time pressure they experience in their job. Speaker A describes
the case where he is busy doing a job, when his boss calls and demands that he
immediately puts up the results of an examination, even though it is late in the
day and A does not see the urgency. The SoC in S1 is a third person, A's boss. The
effect of *want* is that we stay within the perspective of this third person. If we re-
place *want* by *omdat*, as in (16), we need to add a perspective marker like *hij vindt*
('he feels'). This is a clear case of Free Indirect Speech: it cannot be speaker A who

takes responsibility for the content of S2, because the boss and speaker A have such a different view on the issue.

(16) [*S1 hij eist dat je dat nu doet*] *OMDAT* [*S2 hij vindt dat dat nu daar moet hangen*]
[S1 He demands that you do it right now] OMDAT [S2 he thinks that it should hang there right now]

The motivation for the use of *want* by speaker A may well be rhetorical: it makes the information less embedded, and presents the scene more in the here-and-now. This increases the livelihood and dramatic effect of the story.

A very different type of violation of the prototypical *want* profile seems to be used for cooperative reasons. Consider example (17), which comes from a chat session between two secondary school children.

(17) A: *hoihoi*
          hihi
     A: *ik ging femke zoeke :P*
          I went to look for femke :P
     B: *haha*
          haha
     C: *want?*
          WANT?
     A: *dan kon ik daar meej prate :P...*
          then I would have been able to talk to her :P...

This example does not follow the profile in that the *want* is uttered by speaker C but connects to an utterance by speaker A. "It seems as if C is trying to empathize with A, as if the SoC changes from C to A. The effect is that C prompts A to 'say more than [she has]' " (see Lerner [2004: 162–163] for an analysis of such prompts to elaborate).[2] Note that this use of *want* seems an effective strategy to invite the conversational partner to co-construct the utterance. The effect of C using *want?* is that she signals to be willing to stay within A's perspective, as if the SoC does not change between utterances. This mechanism seems to be especially suitable for cases where the speaker is the SoC, and the communicative situation is clearly connected to the current here-and-now. We propose the label "Maintain Perspective" for this strategic use of *want*.

---

2 We thank an anonymous reviewer for bringing this to our attention.

Now consider example (18), a constructed example.

(18) A: John decided to move to the UK
     B: ### WANT?
     A: He doesn't like the political climate in the Netherlands.

This use of *want* strikes us as extremely odd. In (18) the SoC is a third person, *John*, and the causal connection described is to be located in the past. This is not a suitable context for the use of *want* because it does not continue on a first-person perspective located in the here-and-now. It does not come as a surprise, then, that in our corpora we do not find such cases.

Speaker A in (17) could have constructed a completely individual turn, which would have looked like "I went to look for Femke WANT then I could talk to her". Speakers A and C could also have chosen for a different type of co-constructed sequence. Since in Dutch "*Waarom? Omdat...*" is a natural sequence,[3] a standard way of asking for A's reason in Dutch is a division of labor like the following: "A: I went to look for Femke C: Why? A: OMDAT then I could talk to her". This sequence is what speaker C seems to have in mind in the following fragment.

(19)
1   A: *jij bent daan niet*
        you're not daan
2   B: *hoezo?*
        what do you mean?
3   B: *doe niet slim*
        don't be smart
4   A: *jawel*
        yes
5   A: *je bent daan niet*
        you're not daan
6   B: *want?*
        WANT?
7   A: *wats de leipste gam eop de ps2*
        what's the coolest game on the p[lay]s[tation]2
8   A: *?*
        ?
9   C: *omdat je niet als daan doet:p*
        OMDAT you don't act like daan:p

---

3 Whereas "*Waarom? Want...*" is impossible.

In utterance 6 speaker B is following the "Maintain Perspective strategy" identi-fied earlier. He uses *want* to maintain A's perspective. A does not answer, but asks about Playstation 2.[4] Speaker C seems to interpret B's question in 6 as a *why*-question, i.e., as asking for a reason, which she gives in utterance 9, using *omdat*. It is tempting to speculate that C uses *omdat* because as a third person she cannot possibly cooperate with B's use of the "Maintain Perspective" strategy as she in-trudes in the exchange between A and B.

One would expect to find uses of "Maintain Perspective" *want* like (17) and (19) in all forms of informal settings, but we only encountered them in our chat corpus.

Let us summarize the usage characteristics of non-prototypical *want*-cases which seem to express objective relations. Contrary to the default use of *want*, which expresses the subjectivity of the author or first-person SoC, the non-prototypical cases often have third-person SoCs; a typical context for *omdat*. However, we can understand why *want* is used: it provides insight into the mental reasoning of the third-person SoC, as we just demonstrated with the examples (12) and (14). As a result, *want* allows for identification with that SoC. These cases resemble Free Indirect Speech, in which *want* enables identification with a par-ticipant that is not the speaker or author. The same identification effect occurs in cases that we labeled "Maintain Perspective," where *want* (in contrast to *omdat*) allows perspective maintenance of another character, as in (14). This use of *want* was especially found in dialogues (15), (17), (19).

What all these cases have in common, then, is that they make use of *want*'s characteristic to enable identification with a third-person SoC, who is a character in the discourse. Contrary to the default use, this is not identification with a first-person or speaker/author SoC.

6 Discussion and conclusions

We have sketched a profile for the interpretation of the connectives *omdat* and *want* based on corpus analyses of large numbers of occurrences in different me-dia and genres. We noted that there are many deviations from the prototypical use of *omdat* and *want*. Instances of those deviations were analyzed in order to understand the deviating use. We conclude that those deviations can also be interpreted in terms of core elements of the prototypical use. That is why we

---

4 Utterance 7 demonstrates a typicality of chat communication: speakers set up different topic lines, partly because the system has an inherent delay. Here A introduces a new topic in 7 (hip games on the Playstation 2).

conclude that the semantics of *want* and *omdat* should not be considered as hard-wired all-or-nothing rules but that it has a prototype structure, with a core meaning/use, and more peripheral uses. The non-prototypical, peripheral uses are motivated deviations: we need the elements in the profile to understand the deviations.

This does not mean that there are no restrictions on the variation of use. Our analysis would be falsified if there are naturally occurring instances of *omdat* and *want* that cannot be motivated in terms of the prototypical profile. To date we have not found such instances.

Such patterns of clear preferences without black-and-white distinctions are not limited to causal connectives. Prototypicality has been proposed in many cognitive linguistic analyses of various linguistics phenomena (Lakoff 1987), and for causal connectives, it is not limited to Dutch. The prototype structure of causal connectives has been identified across several languages. French, German, and Dutch all have causal connectives that are generally characterized as objective (*parce que, weil, omdat*), and connectives that are characterized as subjective (*car*, *denn,* and *want*). Also, the occurrence of non-prototypical uses has been established for these European languages (Stukker and Sanders 2012): French, German, and Dutch show that, under specific circumstances, causal connectives can be used in non-prototypical contexts.

The non-prototypical use of connectives seems to be systematic. Stukker and Sanders note that this type of usage reflects characteristics which correspond to the connectives' more prototypical use. An illustrative case is that of German *denn*, a subjective connective that sometimes is used in an objective context. When *denn* is used in an objective context, that context contains a first-person SoC more often than when its objective counterpart *weil* is used in such an objective causal context. As first-person SoCs are indications of a high degree of subjectivity, this finding suggests that German objective causal relations marked with *denn* still contain characteristics of *denn*'s more prototypical contexts of use in subjective causal relations (Stukker and Sanders 2012). This shows that non-prototypically used *denn* has more in common with prototypical *denn* than with *weil*.

Such results imply that the organization of lexica of causal connectives reflects linguistic categorization. More specifically, the categories of objective and subjective causality function as the connectives' prototypical contexts of use, whereas the apparent "counterexamples" can be interpreted as less prototypical usage contexts.

Such a take on connectives can be put into the context of the individual's linguistic system as fundamentally grounded in "usage events" (Bybee 2006). In that view, the language system consists of generalizations over individual usage

events, which in turn categorize or license other usage events. This process is often compared to the more general psychological process that the occurrence of any (psychological) event leaves some kind of trace that facilitates their reoccurrence. Because the language system is largely experience driven, frequency of instances is a prime mechanism in its structure and operation. Frequently encountered patterns become "entrenched" in memory; they acquire the status of cognitive routines that are retrieved and applied without requiring conscious attention.

If we look upon the semantics of causal connectives as such cognitive routines, we predict that prototypical usage is more entrenched and hence more frequent. More precisely, Stukker and Sanders (2012) predicted that across languages, connectives specializing in *ob*jective causal relations are used in *sub*jective causal relations *less frequently* than in *ob*jective causal relations. By the same token, connectives specializing in *sub*jective causal relations are used in *ob*jective causal relations *less frequently* than in *sub*jective causal relations. Generally speaking, these hypotheses were corroborated by Stukker and Sanders.

A prototype account of causal connectives might also explain the patterns of usage of these connectives in different genres. Interesting connections can be expected with genre-specific properties, such as the communicative intentions of the speaker/writer. In persuasive genres subjective relations should be more dominant than objective relations (Sanders 1997). Such results raise interesting questions: Do objective connectives occur less in persuasive genres? If connections in persuasive genres are by default subjective, does *want* function similarly in persuasive and informative genres?

One example is the often distantiating use of non-prototypical *omdat* in subjective contexts, identified in Section 4. We found a substantial number of subjective relations in persuasive contexts that are expressed by *omdat*. A case in point is example (20), from an opinionating text (a letter to the editor).

(20) *De algemene conclusie van het artikel luidt dat* [*S1 deze vorm van onderwijs voor kansarme kinderen verpietert*], *OMDAT* [*S2 de stadsdelen ooit verworven bevoegdheden ongaarne afstaan*].
The general conclusion of the article is [S1 that this form of education for underprivileged children is withering] OMDAT [S2 the precincts are reluctant to give up powers obtained earlier].

The author could have used *want* as a connective, especially because S1 expresses a judgment. However, if *omdat* is replaced by *want*, the implication is that the author of the letter agrees with and accepts responsibility for this argumentation. The remainder of the letter makes it clear that the author disagrees

with the argumentation and in fact distances himself from it. In this case *omdat* helps to identify who is responsible for the argumentation. A substantial number of uses of *omdat* in persuasive contexts is used for this type of layered argumentation.

In sum, a detailed analysis of apparent exceptions to the rules of using causal connectives provides much insight into those rules. What is more, such an analysis corroborates those rules, especially in concord with a general pattern of use that is based on more quantitative corpus studies. As such, our work can be seen as a plea for a converging evidence methodology to understanding naturally occurring discourse.

References

Bakhtin, M. 1981. *The dialogic imagination*. Austin: University of Texas Press.

Banfield, Ann. 1982. *Unspeakable sentences: Narration and representation in the language of fiction*. Boston: Routledge & Kegan Paul.

Bybee, Joan. 2006. From usage to grammar: The mind's response to repetition. *Language* 82. 711–733.

Chafe, Wallace. 1994. *Discourse, consciousness, and time. The flow and displacement of conscious experience in speaking and writing*. Chicago: The University of Chicago Press.

Clark, Herbert H. 1996. *Using language*. Cambridge: Cambridge University Press.

De Vries, Jan W. 1971. Want en omdat. *De Nieuwe Taalgids* 64. 414–420.

D-COI. 2006. Dutch Language Corpus Initiative. http://www.utwente.nl/ctit/research/projects/concluded/national/stevin/dcoi.doc/ (accessed April 25, 2013)

Ducrot, Oswald. 1984. *Le dire et le dit*. Paris: Minuit.

Grondelaers, Stefan, Katrien Deygers, Hilde van Aken, Vicky van den Heede & Dirk Speelman. 2000. Het ConDiv-corpus geschreven Nederlands. *Nederlandse Taalkunde* 5. 356–363.

Günthner, Susanne. 1993. ". . . weil – man kann es ja wissenschaftlich untersuchen" – Diskurspragmatische Aspekte der Wortstellung in WEIL-Sätzen. *Linguistische Berichte* 143. 37–55.

Kamalski, Judith, Leo Lentz, Ted Sanders & Rolf Zwaan. 2008. The forewarning effect of coherence markers in persuasive discourse. Evidence from persuasion and processing. *Discourse Processes* 45. 545–579.

Keller, Rudi. 1995. The epistemic *weil*. In Dieter Stein & Susan Wright (eds.), *Subjectivity and subjectivisation*, 16–30. Cambridge: Cambridge University Press.

Knott, Alistair & Robert Dale. 1994. Using linguistic phenomena to motivate a set of coherence relations. *Discourse Processes* 18. 35–62.

Lakoff, George. 1987. *Women, fire and dangerous things: What categories reveal about the mind*. Chicago: University of Chicago Press.

Langacker, Ronald W. 1990. Subjectification. *Cognitive Linguistics* 1. 5–38.

Lerner, Gene H. 2004. On the place of linguistic resources in the organization of talk-in-interaction: Grammar as action in prompting a speaker to elaborate. *Research on Language and Social Interaction* 37. 151–184.

Oostdijk, Nelleke. 2000. The Spoken Dutch Corpus Project. *The ELRA Newsletter* 5. 4–8.

Pander Maat, Henk & Ted Sanders. 2000. Domains of use or subjectivity: The distribution of three Dutch causal connectives explained. In Elizabeth Couper-Kuhlen & Bernd Kortmann (eds.), *Cause, condition, concession, and contrast: Cognitive and discourse perspectives*, 57–82. Berlin & New York: Mouton de Gruyter.

Sanders, José. 2010. Intertwined voices. Journalists' representation modes of source information in journalistic subgenres. *English Text Construction* 3. 226–249.

Sanders, José, Ted Sanders & Eve Sweetser. 2012. Responsible subjects and discourse causality. How mental spaces and perspective help identifying subjectivity in Dutch backward causal connectives. *Journal of Pragmatics* 44. 169–190.

Sanders, Ted. 1997. Semantic and pragmatic sources of coherence: On the categorization of coherence relations in context. *Discourse Processes* 24. 119–147.

Sanders, Ted & Wilbert Spooren. 2009. Causal categories in discourse – Converging evidence from language use. In Ted Sanders & Eve Sweetser (eds.), *Causal categories in discourse and cognition*, 205–246. Berlin & New York: Mouton de Gruyter.

Sanders, Ted & Wilbert Spooren. Submitted. *Causality and subjectivity in discourse: The meaning and use of causal connectives in spontaneous conversation, chat interactions and written text genres*.

Sanders, Ted & Eve Sweetser (eds.). 2009. *Causal categories in discourse and cognition*. Berlin & New York: Mouton de Gruyter.

Sanders, Ted J. M., Wilbert P. M. S. Spooren & Leo G. M. Noordman. 1992. Toward a taxonomy of coherence relations. *Discourse Processes* 15. 1–35.

Spooren, Wilbert P. M. S. & Liesbeth Degand. 2010. Coding coherence relations: Reliability and validity. *Corpus Linguistics and Linguistic Theory 6*. 241–266.

Spooren, Wilbert, Ted Sanders, Mike Huiskes & Liesbeth Degand. 2010. Subjectivity and causality: A corpus study of spoken language. In Sally Rice & John Newman (eds.), *Empirical and experimental methods in cognitive/functional research,* 241–255. Chicago: CSLI/University of Chicago Press.

Stukker, Ninke & Ted Sanders. 2012. Subjectivity and prototype structure in causal connectives: A cross-linguistic perspective. *Journal of Pragmatics* 44(2). 191–231.

Stukker, Ninke, Ted Sanders & Arie Verhagen. 2009. Categories of subjectivity in Dutch causal connectives: A usage-based analysis. In Ted Sanders & Eve Sweetser (eds.), *Causal categories in discourse and cognition*, 119–171. Berlin & New York: Mouton de Gruyter.

Sweetser, Eve E. 1990. *From etymology to pragmatics. Metaphorical and cultural aspects of semantic structure*. Cambridge: Cambridge University Press.

Traugott, Elizabeth C. 1989. On the rise of epistemic meanings in English: An example of subjectification in semantic change. *Language* 57. 33–65.

Traugott, Elizabeth C. 1995. Subjectification in grammaticalization. In Dieter Stein & Susan Wright (eds.), *Subjectivity and subjectivisation*, 31–54. Cambridge: Cambridge University Press.

Van den Hoven, Paul J. 1997. Niet wij doen het, maar het recht! *Taalbeheersing* 19. 207–214.

Verhagen, Arie. 2005. *Constructions of intersubjectivity. Discourse, syntax, and cognition*. Oxford: Oxford University Press.

Wegener, Heide. 2000. *Da, denn* und *weil* – der Kampf der Konjunktionen. Zur Grammatikalisierung im kausalen Bereich. In R. Thieroff et al. (eds.), *Deutsche Grammatik in Theorie und Praxis*. Tübingen: Niemeyer.

Zufferey, Sandrine. 2010. *Lexical pragmatics and theory of mind. The acquisition of connectives as a window on cognitive development*. Amsterdam & Philadelphia: John Benjamins.

Bionotes

Ted Sanders is full professor of discourse studies and Dutch language use in the Utrecht Institute of Linguistics. His work concentrates on discourse coherence, the way in which this is marked (cross-)linguistically, through connectives and other discourse markers. He combines this linguistic work with studies of the cognitive processes of discourse production and interpretation, as well as of acquisition: how do children learn to build discourse? He has (co-)published on these topics in many international journals and book volumes. Address for correspondence: Utrecht Institute of Linguistics, Universiteit Utrecht, Trans 10, Nl- 3512 JK Utrecht, The Netherlands ⟨T.J.M.Sanders@uu.nl⟩.

Wilbert Spooren is full professor of discourse studies and Dutch language use at the Faculty of Arts, Radboud University Nijmegen. He is a member of the Centre for Language Studies. He has published on issues of discourse coherence, subjectivity, and text quality. Address for correspondence: Centre for Language Studies, Radboud University, Nijmegen, The Netherlands ⟨W.Spooren@let.ru.nl⟩.