

The linguistic marking of coherence relations: Interactions between connectives and segment-internal elements

Jet Hoek,¹ Sandrine Zufferey,² Jacqueline Evers-Vermeul³ and Ted J. M. Sanders³

¹ University of Edinburgh, ² University of Bern, ³ Utrecht University, Utrecht Institute of Linguistics OTS

Abstract

Connectives and cue phrases are the most prototypical linguistic elements that signal coherence relations, but by limiting our attention to connectives, we are likely missing out on important other cues readers and listeners use when establishing coherence relations. However, defining the role of other types of linguistic elements in the signaling of coherence relation is not straightforward, and it is also not obvious why and how non-connective elements function as signals for coherence relations. In this paper, we aim to develop a systematic way of categorizing segment-internal elements as signals of coherence relations on the basis of a literature review and evidence from parallel corpora. We propose a three-way distinction between *division of labor*, *agreement*, and *general collocation* to categorize the different ways in which elements inside discourse segments interact with connectives in the marking of coherence relations. In each type of interaction, segment-internal elements can function as signals for coherence relations, but the mechanism behind it is slightly different for each type.

Keywords: Coherence relations, connectives, signals, linguistic marking, parallel corpus

1 Introduction

When readers or listeners are presented with a text, they do not treat the individual clauses and sentences in that text as independent and unrelated. Instead, they try to relate each part of the text, or each *discourse segment*, to the rest of the discourse. By doing so they, ideally, create a coherent mental representation of the discourse. Discourse segments can be related to each other in different ways, for instance through a causal relation, a contrastive relation, or a conditional relation, so language users have to figure out the exact way in which a segment has to be related to another segment. The process of arriving at an appropriate type of coherence relation is often, but not always, facilitated by the presence of explicit linguistic cues in the discourse. The most obvious markers of coherence relations are discourse connectives (e.g., *before*, *if*) and cue phrases (e.g., *for this reason*, *by contrast*). Coherence relations that are marked by a connective or a cue phrase, as in (1a), are commonly referred to as *explicit* coherence relations; coherence relations that are not accompanied by a connective or a cue phrase, as in (1b) are traditionally labeled *implicit* coherence relations. Even

though (1b) lacks the explicit instruction to establish a causal relation that *because* provides in (1a), the relation between the two discourse segments is still easily inferable.

(1a) [Trey pushed Tara,]_{S1} **because** [she threw his baseball bat into the water.]_{S2}

(1b) [Trey pushed Tara.]_{S1} Ø [She threw his baseball bat into the water.]_{S2}

Although the distinction between explicit and implicit coherence relations seems very straightforward, it is not without its problems. Connectives can, for instance, signal a relation that is less specific than the relation that is constructed by language users. (1c), for example, is marked by the temporal connective *after*, but it is very likely that a causal relation will still be inferred between its segments. While both (1a) and (1c) are marked by a connective, they thus seem to differ in the extent to which the relation is explicitly signaled.

(1c) [Trey pushed Tara]_{S1} **after** [she threw his baseball bat into the water.]_{S2}

Similarly, relations that do not contain a connective or a cue phrase can vary in the extent to which the relation remains linguistically implicit. The relation in (2), like the relation in (1b), does not contain a connective, but (2) seems to be crucially different from (1b) in that the relation contains other linguistic cues that help in constructing the relation: the semantic opposites *great* and *horrible*. (2) therefore appears to be less implicit than (1b).

(2) [Jack is a great kid.]_{S1} Ø [His sister is horrible.]_{S2}

When linguistic cues other than connectives or cue phrases are taken into account, the distinction between implicit and explicit relations becomes more fuzzy, and the category of implicit coherence relations becomes much smaller.

1.1 The marking of coherence relations by connectives

Connectives and cue phrases (from now on referred to as *connectives*) provide processing instructions about the way in which two discourse segments should be related to each other (e.g., Sanders & Spooren 2007). The general principle behind the marking of coherence relations seems to be that if the appropriate relation can be constructed without a connective or with an underspecified connective, it can be left implicit or underspecified; if not, the relation should be explicitly marked. This idea is compatible with several different theories of language production. Grice's (1975) Quantity Maxim, for instance, states that a contribution should neither say more nor less than necessary. This two-part maxim has been reformulated by Horn (1984: 13) in terms of speaker-based and hearer-based

principles: the Q principle, which states that you have to “make your contribution sufficient”, and the R principle, which states that you should “make your contribution necessary”. While the Q principle reduces the hearer’s effort to interpret an utterance, the R principle prevents the speaker from wasting effort producing superfluous linguistic content. A coherence relation should therefore be sufficiently marked so that the hearer will be able to construct the appropriate relation, but not be overly or unnecessarily marked so as to limit the speaker’s efforts.

The same idea can also be thought of in terms of in effort versus effect, key notions from Relevance Theory (Sperber & Wilson 1985, Wilson & Sperber 2005); if inferring a specific type of relation requires too much effort for the resulting cognitive effect, another, easier relation will be inferred. Explicitly marking a relation that would otherwise be misinterpreted ensures that the right relation is constructed by the hearer. From a speaker perspective, it costs effect to add a connective to a relation that could already be inferred without it, but the resulting effect is minimal.

Finally, the *Uniform Information Density* (UID) hypothesis (Frank & Jaeger 2008, Levy & Jaeger 2007) proposes that speakers “structure their utterances so as to avoid peaks or troughs in information density” (Levy & Jaeger 2007: 1). If a coherence relation can be established without the presence of a connective or cue phrase, inserting a connective or cue phrase would lead to a trough in information density, since it does not add any extra information to the utterance. Conversely, leaving a relation that is hard to interpret without a connective results in a peak in information density, since too much information has to be extracted from the other linguistic elements in the utterance.

All these accounts predict that speakers use a connective when it contributes essential information to the discourse. This idea is also in line with a ‘collaborative’ view of communication, under which speakers or readers can be assumed to use connectives (or cohesive devices in general) to successfully transmit their message, and text recipients can be assumed to pay attention to these signals so as to successfully understand the message (e.g., Tanskanen 2006). All these accounts suggest that if the connective is barely informative or even entirely redundant, speakers will be more inclined to leave it out, in which case the relation will be implicit, or use a more general connective, in which case the relation will be underspecified. If another element within a discourse segment already signals or partly signals how that segment should be related to another segment from the discourse, this would eliminate or reduce the amount of information a connective would contribute. Other linguistic elements that convey information or raise expectations about the type of coherence relation that should be constructed are thus expected to influence the marking of coherence relations by connectives.

1.2 The marking of coherence relations by other linguistic elements

Research on the marking of coherence relations has mostly been focused on connectives and cue phrases, because these are the only linguistic elements that by definition express relational meaning;

they signal the meaning that the relation adds to the meaning of the segments in isolation. As such, connectives and cue phrases have a special status in cognitive and linguistic studies of coherence relations, especially because there are systematic restrictions on their meaning and use and because these systems organize the lexica of connectives in various languages (e.g., Knott & Dale 1994, Knott & Sanders 1998). Connectives and cue phrases being the only elements that exclusively express relational meaning does not necessarily mean, however, that connectives and cue phrases are the *only* indicators for coherence relations. By limiting our attention to connectives, we are therefore likely missing out on important other cues readers and listeners use when establishing coherence relations.

The most elaborate research effort to identify other signals for coherence relations has been the recently released *RST Signalling Corpus* (Das, Taboada & McFetridge 2015; see Das & Taboada 2013, 2018, and Taboada & Das 2013 for analyses based on the RST Signalling Corpus), in which linguistic cues that signal coherence relations annotated in the *RST Treebank* (Carlson, Okurowski & Marcu 2002) are identified. However, since linguistic elements other than connectives do not directly signal coherence relations, it is not obvious how to group or categorize signals. For example, Taboada and Das (2013: 258) indicate that the numerical element *five* in the first segment of (3) is a signal for the ELABORATION relation between S_1 and S_2 , with the five names in S_2 specifying the contents of the “five”.¹ The relation in (3) is then categorized as being signaled by a *numerical* signal.

- (3) [This maker of electronic devices said it replaced all **five** incumbent directors at a special meeting...]S₁ [Elected as directors were Mr. Hollander, Frederick Ezekiel, Frederick Ross, Arthur B. Crozier and Rose Pothier.]S₂

Although it seems indeed plausible that *five* in S_1 and the list of five names in S_2 are in this case important linguistic elements when inferring a coherence relation between these two segments, the *numerical* signal does not in itself signal an ELABORATION relation. *Five*, or any other number, may just as well occur in any other type of relation. In (4), the number *five* is for instance in contrast with the number *twelve*, and could as such be taken to signal a CONTRAST relation.

- (4) [Jane packed five pairs of socks for her school trip.]S₁ [Frank brought twelve pairs.]S₂

Instead of considering both (3) and (4) as instances of *numerical* signaling, they could receive more specific labels. (3) could for instance be categorized as *numerical-lexical enumeration* and (4) as a *numerical pair*, which would allow us to distinguish between these two relations, while at the same type attributing signal status to other parts of the relation as well (the list of names and *twelve*, respectively). However, other examples may provide new ideas for labels and groupings. In sum,

¹ It could, however, be argued that the set of five in S_1 is a different set than the five names in S_2 .

determining which parts of a relation function as a signal is not straightforward, nor is it obvious how to categorize the identified signals.

While the *RST Signalling Corpus* is a valuable inventory of potential signals, both within the segments and outside them (e.g., text structure, genre), it does not link signals to relation types in a systematic way and does not comment on *how* or *why* the indicated signals function as cues for coherence relations. In addition, since the annotation was mostly focused on relations without connectives, the *RST Signalling Corpus* does not identify potential additional signals in relations that contain a connective. The current paper explores the marking of coherence relations by connectives on the one hand, and other types of cues on the other. Specifically, we will investigate how linguistic elements within the segments of a coherence relation, i.e., *segment-internal elements*, can contribute to the marking of the relation, and how the presence of other signals relates to the presence of connectives.

Connectives encode instructions on how to relate discourse segments to each other, which is why they themselves are not part of the segments. Throughout this paper, connectives will therefore be left out of the segmentation of the relations whenever possible. Often, the marker will appear either before S_1 or between S_1 and S_2 (i.e., at the head of S_2), but this is by no means a categorical feature of connectives, since they can also appear in clause-medial or clause-final position (in S_2). The class of connectives includes several different grammatical categories, such as conjunctions, adverbs, and idiom chunks. While conjunctions are restricted to appear either between the two segments or before S_1 , adverbs and idiomatic phrases can also appear in clause-medial or clause-final position. This is a crucial distinction, since even if connectives appear in clause-medial position, this does not make them part of the segment, or segment-specific cues. We consider elements to be segment-internal if they are integrated in and are part of the propositional content of the clauses that are, or are part of, the segments of a coherence relation.

1.3 The interaction between connectives and segment-internal elements

There are several segment-internal features that have been linked to particular types of coherence relations. These segment-specific elements include a wide range of linguistic categories, such as complex phrases, lexical items, modal markers, and verbal inflection. The features can either occur in one of the segments or in both of the segments. However, it seems that not all linguistic elements that have been associated with a specific type of coherence relation signal the relation in the same way, and there appear to be differences in the way in which the presence of a specific linguistic element in one of the segments of a relation can impact the marking of that relation by means of a connective. In this paper, we argue that there are three distinct ways in which segment-internal elements systematically interact with the connective that marks a coherence relation, which we label *division of*

labor, *agreement*, and *general collocation*.² In all three types of interactions, the segment-internal element can function as a cue for which type of coherence relation should be constructed or expected, but the reason why the segment-internal element functions as a signal for a coherence relation is slightly different in each interaction. The three-way categorization is based on specific linguistic elements that have been reported to often co-occur with specific types of connectives or coherence relations in monolingual corpus studies or experimental studies; the existence of the three types of interactions is demonstrated using data from a parallel corpus. Section 2 first introduces the parallel corpus used in this study and the rationale behind using translations to study the marking of coherence relations. Sections 3, 4, and 5 discuss, respectively, *division of labor*, *agreement*, and *general collocation*.

2 The marking of coherence relations in parallel corpora

The current paper makes use of parallel corpus data to investigate the marking of coherence relations by connectives and cue phrases on the one hand and segment-internal elements on the other. Section 2.1 explains how translations can provide insights into which parts of a text fragment contribute to the marking of coherence relations and introduces the specific discourse-annotated parallel corpus used in this paper. Section 2.2 gives a brief overview of the annotation model used to annotate the coherence relations in the parallel corpus.

2.1 Coherence relations in translation

Monolingual corpora are extremely valuable resources for language research, but when studying *meaning* using a monolingual corpus, researchers still mainly have to rely on their own interpretations, since “meaning is not directly observable” (Noël 2003: 758). When it comes to the interaction between segment-internal elements and connectives, it is not necessarily obvious what and how each element contributes to the overall interpretation of a relation. A proposal for an alternative method to research meaning is to make use of parallel corpora, which consist of a source text (ST)

² *Collocation* is a term widely used in corpus linguistics to refer to “words that occur (with a frequency that is significantly above chance-level) in a given span around the node word” (Gries & Stefanowitsch 2004:100), with the words that often co-occur together being referred to as *collocates*. While we do not use collocational analysis in this paper, all sets of connectives/coherence relations and segment-internal elements we refer to have been shown or argued to often occur in each other’s vicinity, which is why we adopted the term *collocation*. Technically, division of labor and agreement interactions could also be considered examples of collocations, but with properties that – we argue – distinguish them from the type of interaction we labelled *general collocation*, see Sections 3 and 4.

and one or multiple translations (target texts: TTs) (e.g., Dyvik 1998, Melamed 2001, Noël 2003, Teubert 1999). In this approach, the translator is treated as a naive ‘annotator,’ whose main purpose was to accurately convey the meaning of the ST in the TT. Variety in the linguistic means used to arrive at a similar meaning can inform researchers about the contribution of individual elements or constructions to the meaning of a text fragment.

Connectives are known to be volatile items in translation (Halverson 2004, Zufferey & Cartoni 2014). Translations often use an equivalent connective to translate the original connective, as in (5), but can also leave out the connective entirely, as in (6), rephrase the meaning of the connective using different linguistic means, as in (7), or use a connective that does not express the same meaning as the original connective, as in (8).

- (5) EN [This is nothing to do with aeroplanes,] **although** [I could speak about flights if you wished me to.]
 ES [Esto no tiene nada que ver con aviones,] **aunque**, [si lo desean, puedo hablar de volar.]
- (6) EN [I would like a clear answer from the Commission as to why it is failing to do what This Parliament has asked it to do and what it was instructed to do by the Council at The spring summit,] **because** [this is not acceptable.] {ep-02-11-20}
 DE [Ich bitte die Kommission um eine klare Antwort, warum sie der Forderung des Parlaments und der Anweisung des Rates, die ihr auf dem Gipfeltreffen im Frühjahr erteilt wurde, nicht nachkommt.] Ø [Dies ist nicht hinnehmbar.]
- (7) EN [Currently there are no equivalent tests,] **so** [those tests are for the moment standard.] {ep-02-06-13}
 FR [Il n'existe pas à l'heure actuelle de tests équivalents,] [**ce qui leur confère la qualité de tests standards.**]
 ‘Currently there are no equivalent tests, which gives them the status of standard tests.’
- (8) EN [I think honourable Members are aware at least of the organizational complexity] **because** [all this has major practical and legal implications.] {ep-96-11-19}
 NL [Dit brengt praktisch en juridisch gezien heel wat met zich mee,] **en** [ik denk dat de geachte afgevaardigden wel inzien hoe moeilijk dit te organiseren is.]
 ‘This has major practical and legal implications, **and** I think the honourable Members are aware of how difficult this is to organize.’

The translations in (5)–(8) show how coherence relations can vary in the extent to which they are explicitly signaled, with the translated relations in (6) and (8) being less explicit than the English originals, and the French translation in (7) being arguably even more explicit than the original English

relation. The variations introduced in the linguistic encoding of the marking of coherence relations in translation supposedly have no detrimental effects on the meaning of a text as a whole. This makes translation data especially suited to research the different linguistic resources that can be used to signal coherence relations. The potential of using parallel corpora to research discourse coherence phenomena has already been demonstrated in several other studies (e.g., Cartoni, Zufferey & Meyer 2013, Cartoni, Zufferey, Meyer & Popescu-Belis 2011, Hansen-Schirra, Neumann & Steiner 2007, 2012, Levshina & Degand 2017).

In the current paper, we make use of a multilingual discourse-annotated parallel corpus based on the *Europarl Direct* corpus (Cartoni, Zufferey & Meyer 2013, Koehn 2005) to supply additional evidence for the existence of three different types of interactions between connectives and segment-internal elements. The corpus consists of approximately 2000 explicit English coherence relations (with the connectives *also*, *although*, *because*, *but*, *if*, *in addition*, *so*, *unless*) with translations into Dutch, German, French, and Spanish. All English relations were annotated using CCR (*Cognitive approach to Coherence Relations*, see Section 2.2). We hypothesize that the different types of interactions between segment-specific elements and connectives can be distinguished in translation and formulate specific expectations for the translation of each type of interaction in the three corpus data sections (Sections 3.2, 4.2, and 5.2).

2.2 The *Cognitive approach to Coherence Relations*

The coherence relations in the English part of the parallel corpus used in the paper have all been annotated using the *Cognitive approach of Coherence Relations* (CCR; Sanders, Spooren & Noordman 1992 and later work). This framework uses a set of primitives, rather than single end labels to classify coherence relations. Each basic primitive has at least two values: POLARITY (positive, negative), BASIC OPERATION (causal, conditional, additive), SOURCE OF COHERENCE (objective, subjective), ORDER OF THE SEGMENTS (basic, non-basic), and TEMPORALITY (temporal, non-temporal) (Evers-Vermeul, Hoek & Scholman 2017, Sanders et al. 1992).³ In addition to the basic primitives, several other features have been identified that, unlike the four main primitives, apply only to a small subset of relations, e.g., VOLITIONALITY and PURPOSE (see Hoek, Evers-Vermeul & Sanders *submitted*).

CCR considers coherence relations to be cognitive constructs; they are properties of the mental representation of a text, rather than of the linguistic realization of a text. However, connectives constitute linguistic cues as to which relation has to be inferred. The distinction between the

³ See Sanders et al. (1992) for the original CCR taxonomy, and Hoek, Evers-Vermeul, and Sanders (*submitted*) for a more elaborate overview of the considerations and decisions made while annotating the English ST relations in the dataset used in the current study.

coherence relation on the one hand and the linguistic signals for that relation on the other can for instance also be found in Halliday and Hasan (1976; *coherence* versus *cohesion*) and Renkema (2009; *adjunction* versus *conjunction*).

While CCR and its primitives were originally developed to depict coherence relations, they can also be used to describe what is signaled by connectives (e.g., Knott & Sanders 1998). Connectives provide language users with instructions on how to relate two (or, occasionally, multiple) discourse segments to each other, but not all connectives are equally specific.⁴ *Because*, for instance, signals a positive value for POLARITY and a causal value for BASIC OPERATION, while the only thing that *but* signals is a negative value for POLARITY. As the relation in (1c), repeated below as (9), illustrates, it is possible for a connective to appear in a coherence relation that is more specific than what it in itself signals; *after* signals positive temporal relations, but the coherence relation in (9), i.e., the relation in the mental representation constructed on the basis of (9), is more appropriately labeled a positive causal relation.⁵

- (9) [Trey pushed Tara]_{S1} **after** [she threw his baseball bat into the water.]_{S2}

Since CCR can be used to describe what is signaled by connectives, it should also be applicable to depicting what is signaled by other linguistic elements that contribute to the marking of coherence relations. The semantic opposition in (4), for instance, indicates that the relation has a negative value for POLARITY.

In the remainder of the paper, the CCR primitives will be used to depict the type of coherence relations that hold between two discourse segments, as well as to describe what is explicitly signaled by connectives and other signals of coherence relations. The individual primitives allow us to determine to what extent linguistic cues, connectives or other elements, explicitly signal the relation.

3 Division of labor

⁴ CCR considers coherence relations to be cognitive constructs; they are properties of the mental representation of a text, rather than of the linguistic realization of a text. However, connectives provide linguistic cues as to which relation has to be inferred. A similar distinction is for instance made by Renkema (2009), who distinguishes *conjunction*, the way in which clauses are linguistically related, from *adjunction*, the way in which the information conveyed by clauses is related to each other.

⁵ Causal relations are more specific than temporal relations, see e.g., Evers-Vermeul, Hoek, and Scholman (2017), Sanders et al. (1992), Spooren (1997).

There are several segment-specific cues that appear in one or both of the segments of a coherence relation that seem to eliminate or reduce the need for a connective because the segment-internal elements themselves already signal which type of relation should be inferred between the two discourse segments. In these cases, there appears to be a *division of labor* between the segment-internal cue and the connective; either the connective signals the relation, or an element inside one or both of the segments does so. It even seems possible for the segment-internal element and the connective to each signal part of the relation that should be inferred.

3.1 Division of labor between connectives and segment-internal elements

An obvious example of a *division of labor*-type of interaction between parts of the discourse segments and the connective can be found in relations with what Hoey (1983) labels *lexical signaling*. Hoey (1983: 44–53) argues that the relation in (10) can be, somewhat wordily, paraphrased as (11).

- (10) *I beat off the attack **while** I was on sentry duty, by opening fire **when** I saw the enemy approaching.*
- (11) *The cause of my opening fire was that I saw the enemy approaching. The circumstances of my seeing the enemy approaching was that I was on sentry duty.*

The coherence relations signaled by connectives in (10) are expressed by lexical items in (11). This eliminates the need for connectives, and vice versa; the presence of connectives in (10) renders the use of the lexical signals from (11) obsolete.

The fragment in (11) is a very contrived example. Actual instances of this type of signaling can be found in the *Penn Discourse Treebank* (PDTB; Prasad et al. 2008). In the absence of a connective, the PDTB instructs annotators to insert a connective that best expresses the inferred relation. However, in situations in which inserting a connective may lead to “redundancy in the expression of the relation [...] because the relation is alternatively lexicalized by some ‘non-connective expression’”, annotators do not have to supply a connective (PDTB Research Group 2007: 22). Instead, they are instructed to categorize the marking of the coherence relation as *AltLex* (*Alternative Lexicalizations*). An example of a RESULT relation marked by an alternative lexicalization can be found in (12).

- (12) [*Ms. Bartlett’s previous work, which earned her an international reputation in the non-horticultural art world, often took gardens as its nominal subject.*]_{S1} [*Mayhap this metaphorical connection made the BPC Fine Arts Committee think she had a literal green thumb.*]_{S2} (PDTB Research Group 2007: 22–23)⁶

⁶ We copied the original PDTB segmentation.

In (12), the underlined alternative lexicalization indicates, with the verb *made*, that S_2 provides the result of S_1 , which is similar to what the function of, for instance, *so* would be. Adding *so* would be superfluous or result in a fragment of which the interpretation does not correspond to the current construction, for instance a conclusion. Some of PDTB's *Alternative Lexicalizations* could be considered more elaborate instances of connectives or cue phrases, for example *the most likely reason for this disparity is that* (PDTB Research Group 2007: 23) being a more specific version of *the reason is that*. The crucial differences, however, are that *AltLexes* often include an anaphoric reference to the other segment (PDTB Research Group 2007) and that they often, as in (12), seem to contain propositional content.⁷ As such, many *AltLexes* are more appropriately considered as part of the segments than as the marking of the coherence relation only. *Alternative Lexicalizations* can thus be considered as segment-internal elements that convey information about the coherence relation that should be established, although each instance tends to be novel and can signal a very broad range of relation types.

A more specific segment-internal cue that has been linked to a particular type of coherence relations is *negation*.⁸ Using corpus data from the PDTB, Webber (2013) shows that the presence of negation markers in the first segment is a cue for CHOSEN ALTERNATIVE relations. In CHOSEN ALTERNATIVE relations, which are often signaled by *instead*, the two segments present alternatives, of which one is taken (PDTB Research Group 2007: 36), as in (13).

- (13) [You cannot buy tickets online.]_{S1} **Instead**, [you have to make sure you get to the venue on time.]_{S2}

Webber (2013) argues that a negation marker suffices to indicate that one of the alternatives has been excluded. This does not only explain why negation is a common feature of CHOSEN ALTERNATIVE relations in general, but also why it is even more frequent in CHOSEN ALTERNATIVE relations that are not marked by a connective (Webber 2013). In addition to negation markers, Webber (2013) points toward other elements that create negative assertions (a subset of downward entailing constructions and modal markers that indicate that an event does not hold) as cues heavily associated with CHOSEN ALTERNATIVE relations. These cues are also more frequent in implicit than in explicit CHOSEN ALTERNATIVE relations. Since a negation element in itself signals that only one of the two segments of a relation holds, which is very similar to the function of *instead*, and since negation elements occur

⁷ Connectives are generally considered to have procedural meaning, rather than propositional meaning (e.g., Blakemore 1987, Wilson & Sperber 1993).

⁸ We use *negation* here to refer to linguistic elements that signal negative polarity, including both syntactic and morphological negation.

much more frequently in implicit CHOSEN ALTERNATIVE relations, it seems that in this combination of segment-internal cue and connective, the task of signaling the relation can be taken up by either one of these linguistic elements. As such, negation+*instead* can be considered an example of a *division of labor*-type of interaction.

Another example of an interaction between segment-internal features and connectives that appears to fall into the division of labor category are certain combinations of verb tenses and connectives specifying ORDER (e.g., Kehler 1994, Lascarides & Asher 1993).

(14) [Hank was crying non-stop]_{S1} **after** [his girlfriend broke up with him.]_{S2}

(15) [Hank was crying non-stop.]_{S1} (**after**) [His girlfriend had broken up with him.]_{S2}

In both (14) and (15), S₂ takes place before S₁. In (14), this is signaled by the connective. In (15), the pluperfect in S₂ explicitly places the event before the event in S₁, expressed in the simple past.⁹ While in (15) the connective *after* may be added, it is certainly not necessary to arrive at the same interpretation. Removing the connective from (14), however, changes the interpretation of the relation; in (16) the crying is more plausibly interpreted as preceding and probably even causing the break-up.

(16) [Hank was crying non-stop.]_{S1} [His girlfriend broke up with him.]_{S2}

The combination of verb tenses in (15) signals that the relation has a non-basic order and reduces or even eliminates the need for a connective indicating non-basic order. Conversely, if a connective specifies that S₂ took place before S₁, both segments can be expressed in the same tense. A similar example in which a combination of verb tenses can make a connective superfluous is two segments expressed in the French *passé composé*, which has been claimed to encode sequential, or basic, temporal order, in combination with the connective *puis*, which indicates the same thing. Sentences that have both segments in the *passé composé* and are marked by *puis* are dispreferred over sentences in which the main verb tenses and the connective do not both explicitly encode sequentiality (Grisot & Blochowiak *submitted*).

Other examples of division of labor type of interactions are a semantic opposition between elements in both segments and contrastive connectives, as illustrated in (2), repeated here as (17), and focus markers such as *only* in S₁ and contrastive connectives, as illustrated in (18). *Only* in S₁ indicates that the content of this segment, unlike its alternatives, holds true, which is highly similar to the function that *but* would have in (18) if *only* were left out (Carlson 2014, Umbach 2005). Note that in both these examples, a less specific connective like *and* would work fine as well. Both (17) and (18) are negative

⁹ In both cases, it is likely that a causal relation is inferred on top of the temporal relation.

additive relations. In case the relations are marked by *and*, the connective signals that the BASIC OPERATION of the relation is additive, while the semantic opposition or the focus marker would indicate that the POLARITY of the relation is negative.

- (17) [Jack is a great kid.]_{S1} (**but**) [His sister is horrible.]_{S2}
 (18) [Kathlyn only liked her uncle]_{S1} (**but**) [She didn't like any of her other family members.]_{S2}

In division of labor types of interactions between segment-internal features and connectives, both types of signals are in themselves able to explicitly mark the same primitives of a coherence relation and presence of both types of signals is redundant. This makes it distinctly different from the other two types of interactions, as we will demonstrate in Sections 4 and 5. We discuss the implications that the differences between the three types of interactions have for processing in Section 6.

3.2 Division of labor in translation

In division of labor types of interactions between connectives and segment-internal elements, the presence of the segment-internal feature can make the connective, or part of the connective, redundant, and vice versa. When it comes to translation, it can therefore be expected that such segment-internal features can appear in the translation **instead of** the connective. Alternatively, a segment-internal element can be used to signal part of the primitive values signaled by the original ST connective. In (19), the original English relation uses *also* as a connective to link two arguments together. The German translation does not include a connective, but uses *auch* ‘also’ as an NP modifier on the subject, which is obvious from the word order in *S*₂.¹⁰ Despite the different constructions, the original and the translation yield very similar interpretations. This suggests that, at least in some relations, NP modifiers signaling additivity can function as an explicit signal for additive relations, and that this signal can replace the need for an additive connective.

- (19) EN [Strict time-limits would clearly be very useful here.] **Also**, [the idea of a complaints register accessible on the Internet links the last debate with this one rather nicely and empowers those who have difficulties in this area.] {ep-99-01-13}

DE [Es liegt auf der Hand, daß strikte Zeitvorgaben hier sehr nützlich wären.] [Auch die Idee eines über das Internet zugänglichen Beschwerdereregisters stellt eine recht

¹⁰ With German being a V2 language, the difference between *auch* as an NP modifier on the subject and *auch* as a connective can be clearly distinguished by the position of the verb. If *auch* is used as a connective, the verb appears between *auch* and the NP; if *auch* is an NP modifier, as in (19), the verb appears after *auch* and the rest of the subject.

passende Verbindung zwischen dieser und der vorangegangenen Debatte her und gibt denjenigen ein Mittel an die Hand, die auf diesem Gebiet Schwierigkeiten haben.]

‘It is obvious that strict time-limits would be very useful here. The idea of a complaints register accessible on the Internet as well ...’

Conversely, a connective in the TT can be used to replace a segment-internal element in the ST. In (20), for example, the Dutch connective *zodat* ‘so that’ is a connective signaling that a PURPOSE relation should be inferred between the segments. The English original, however, uses the more general causal connective *because*. The difference between the segments of the ST and the TT is that the English original includes *we want*, which does not show up in *S*₂ of the Dutch translation. Interestingly, the meaning of *we want* exactly expresses the intentionality of the causal relation that distinguishes PURPOSE relations from other types of causal relations (e.g., Reese, Hunter, Asher, Denis & Baldridge 2007). This example thus demonstrates that the combination of *because* and an expression of intentionality inside the antecedent can together signal a PURPOSE relation: an example of division of labor between a segment-internal element and a connective.

(20) EN [*Help us to be more precise with the road-map*] **because** [*we want* *to follow it.*]
{ep-00-04-11}

NL [*Help ons de routebeschrijving nauwkeuriger te maken,*] **zodat** [*we hem kunnen volgen.*]
‘Help us make the road-map more precise, **so that** we can follow it.’

A highly frequent example of division of labor in the corpus can be found in English relations marked by *unless* (negative conditional relations). While all target languages possess a grammaticalized equivalent connective or cue phrase (DE *es sei denn*, ES *a menos que*, *a no ser que*, FR *à moins que*, NL *tenzij*), these are often not used to translate *unless*. Frequently, the translations make use of a connective equivalent to *if* (DE *wenn*, ES *si*, FR *si*, NL *als*). In these translations, a negation element is introduced in one of the segments (usually the antecedent), as in (21). If, however, a connective equivalent to *unless* is used, no negation elements are introduced inside the segments, as is illustrated by (22).

(21) EN **Unless** [*we take that way,*] [*the only alternative will be more misery, more destruction and more death.*] {ep-02-04-09}

FR **Si** [*nous n'empruntons pas cette voie,*] [*nous assisterons à plus de misère, plus de violence et plus de morts.*]

‘**If** we don’t take this path, we will witness more misery, more destruction and more death.’

(22) EN [*There is no reason for their disappearance*] **unless** [*we condemn them to such a fate.*]

{ep-01-12-17}

FR [Il n'y a aucune raison qu'ils disparaissent] **à moins que** [nous les condamnions à un tel sort.]

‘There is no reason they would disappear **unless** we condemn them to such a fate.’

Table 1 provides an overview of how often a negation element is inserted into a TT segment in translations with a connective equivalent to *unless* versus in translations with a connective equivalent to *if*. Negation is only added to a TT segment when the negative value for POLARITY is not expressed by the TT connective, as is the case with *if* ($\chi^2(1)=638.99, p<.001$).¹¹ Since this trade-off pattern was the same in all four language pairs, we grouped all data together.

Table 1. Translations of *unless* into Dutch, German, French, and Spanish

	‘unless’	‘if’
+ negation in TT	0	304
– negation in TT	339	0

A similar division of labor pattern can be found in the translation of English relations marked by *if* into, for instance, German subjunctive mood (*Konjunktiv II*) can be used to express conditionality, as is illustrated by (23).

(23) EN **If** [there had been a check with cards,] [maybe we could have avoided the problem of having a head count.] {ep-00-04-13}

DE [Wäre eine Überprüfung mit Karten vorgenommen worden,] [hätten wir vielleicht das Problem des Abzählens der einzelnen Mitglieder umgehen können.]

‘Had a check with cards been made, we would perhaps have been able to avoid the problem of having a head count.’

Table 2 shows that if the German translation uses a subjunctive, it does not include a connective, and that when the translation uses a positive conditional connective equivalent to *if* (*wenn*), it does not introduce a subjunctive mood to the antecedent ($p<.001$, Fisher’s exact test).

Table 2. Translations of *if* into German

	<i>wenn</i> ‘if’	no connective
+ subjunctive in TT	0	31
– subjunctive in TT	187	0

¹¹ All statistical tests were performed in R (R Core Team 2016, version 3.2.4).

The combinations of segment-internal features and connectives discussed in this subsection are all clear examples of division of labor type of interactions. In translation, a feature can be used to replace the ST connective or part of it, and vice versa. This translation pattern is distinct from what can be seen in agreement and general collocation types of interactions, as will be illustrated by the data presented in the next two sections.

4 Agreement

Even if the meaning signaled by a segment-internal feature overlaps with the meaning signaled by a connective, this does not necessarily mean that the presence of one can make the presence of the other redundant, as is the case with division of labor type of interactions. In these combinations of segment-internal elements and connectives, the presence of the feature does not affect whether or not a connective is used to signal the same meaning. In these cases, there seems to be *agreement* between the feature and the connective.

4.1 Agreement between connectives and segment-internal elements

One example of an agreement type of interaction can be found in non-volitional causal relations in Dutch. Non-volitional causal relations feature a cause that does not involve a volitional agent, and a result, as in (24) (e.g., Mann & Thompson 1988, Stukker, Sanders & Verhagen 2008).

- (24) **Because** [*the airport's main runway was covered in snow*]_{S1} [*all flights were delayed*]_{S2}
 (25) **Doordat** [*er te veel sneeuw op de grootste vertrekbaan van het vliegveld lag*]_{S1} [*hadden alle vluchten vertraging*]_{S2}

(25) is the Dutch equivalent of (24). In this example, the relation is marked by *doordat* ‘because of the fact that’ which is a positive causal connective specified for non-volitionality (Stukker et al. 2008). If the general positive causal connective *omdat* had been used, or even no connective, the relation would still have been interpreted as a non-volitional causal relation, since the presence of the non-volitional event in the antecedent is sufficient to signal the non-volitionality. By definition, all non-volitional causal relations contain a non-volitional event, and yet Dutch has a specific causal connective that appears in these contexts. This type of interaction is thus crucially different from the type of interaction between the segment-internal features and connectives found in the previous section. While both the segment-internal element and the connective signal a similar feature, they tend to appear together. In these cases, there does not seem to be a division of labor between the different signals, but rather agreement.

A similar type of agreement can be found in positive subjective causal relations. Such relations contain a conclusion, claim, or judgment by the speaker, with an accompanying motivation. Many consequents of subjective causal relations feature subjective words that convey the speaker’s attitude (e.g., *stupid*) or otherwise indicate that the content of the segment stems from the speaker’s mind (e.g., *obviously*, modal verbs), as in (26) (e.g., Canestrelli 2013, Wei 2018). However, this need not necessarily be the case, as is illustrated by (27).

- (26) [That guy must obviously be stupid,]_{S1} because [he tried to dry his socks in the microwave.]_{S2}
- (27) [Charlotte is dating someone,]_{S1} because [she has cancelled plans with us three times in the past two weeks.]_{S2}

In isolation, S₁ from (27) could just as well, if not more likely, be a fact. In S₂, however, a motivation is added, indicating that the relation involves the speaker’s reasoning, in which the first segment is treated as a claim. Several languages have been claimed to possess specific subjective causal connectives, which would be used to mark relations like the ones in (26) and (27). Examples of subjective causal connectives are Dutch *want*, German *denn* and *da*, French *car*, and Mandarin Chinese *jiran*, which can all be translated as ‘because’ (e.g., Li, Sanders & Evers-Vermeul 2016, Pit 2003). Similar to non-volitional connectives marking non-volitional causal relations despite the presence of a strong non-volitional cue in the form of an event without an agent, subjective causal connectives are typically used to mark subjective causal relations even if the consequent contains subjective elements. In Dutch, for instance, the most prototypical connective in both (26) and (27) would be *want*; in German, it would be *denn* or *da*.

In Section 3.1, negation elements were discussed in relation to CHOSEN ALTERNATIVE relations, following Webber (2013). Asr and Demberg (2015) also look at the presence of negation markers in coherence relations, but consider a wide range of relation types. Their findings are in line with Webber (2013) when it comes to CHOSEN ALTERNATIVE relations and even show that the presence of a negation marker is most strongly associated with this type of relation. In addition, they find that negation markers also often appear in other types of negative relations, for example in EXPECTATION, CONTRAST, COMPARISON, and CONTRA-EXPECTATION relations (all PDTB 2.0 labels). For most of these relations – all but EXPECTATION – however, it is not the case that negation markers appear more often in the implicit than in the explicit realizations of the relation. Since EXPECTATION, CONTRAST, COMPARISON, and CONTRA-EXPECTATION are all negative relations, the meaning signaled by the segment-internal negation element seems to overlap with the connectives prototypically used to mark these relations (*but, however, although, even though*); both the feature and the connective indicate a negative value for POLARITY. Unlike for CHOSEN ALTERNATIVE relations, the presence of a negation

appears, by itself, not to be sufficient to signal these types of relations.¹² Consider for instance (28) and (29), negative causal relations that when using PDTB 2.0 would be classified as EXPECTATION relations.

- (28) [Gary has never finished a knitting project,]_{S1} **even though** [he loves to knit]_{S2}
 (29) [Gary recently threw out all his knitting supplies,]_{S1} **even though** [he loves to knit]_{S2}

In both (28) and (29), S_2 denies an expectation that S_1 sets up, the underlying assumption being that people who never finish knitting projects or get rid of their knitting supplies do not love knitting. (29) demonstrates that a negation element is not required for negative causal relations, while (28) shows that it is also not sufficient to signal a negative causal relation. It does not even suffice to signal just the POLARITY of the relation, since substituting *even though* with *because* would result in a very different, somewhat incomprehensible fragment.

The interaction between subjective causal and non-volitional causal connectives and their corresponding segment-internal cues, as well as the interaction between negation elements and connectives marking EXPECTATION, CONTRAST and CONTRA-EXPECTATION relations, is thus crucially different from the interactions discussed in Section 4.2. Unlike in division of labor type of interactions, the presence of a segment-internal cue in agreement type of interactions does not seem to make the presence of a connective redundant, and vice versa, even though the two signals overlap in terms of the primitive values they explicitly encode. Although this type of interaction appears to go against what would be predicted on the basis of the linguistic theories discussed in Section 1.1, *agreement* is a very common linguistic phenomenon (e.g., number, gender, or case agreement), and it has often been demonstrated that language tends to be less economical, and thus more redundant, than minimally required (e.g., Bazzanella 2011, Horn 1993). Both observations also seem relevant to the marking of coherence relations; often, a connective agrees with the type of relation it occurs in (i.e., causal connectives usually appear in causal relations, temporal connectives most prototypically occur in temporal relations, etc.), and more coherence relations are explicitly marked than would be absolutely necessary, as is for instance illustrated by the fact that many of the examples in this paper are perfectly acceptable both with and without connective, e.g., (24), (26), and (27). In Section 5, we will discuss a final type of interaction, *general collocation*, in which there is no overlap between the meaning encoded by the connective and the segment-internal cue.

¹² It should be noted that COMPARISON is the most general type of negative relation in the PDTB 2.0 inventory and that CONTRAST includes many more types of relations than, for instance, the ones in (17) or (18). Many relations in the CONTRAST class are, for example, JUXTAPOSITIONS, which do not feature direct opposites, like *nice* and *horrible*, but rather non-identical alternatives, like \$5 and \$10.

4.2 Agreement in translation

As was argued in the previous section, an example of an agreement type of interaction can be found in subjective causal relations in Dutch. In general, the subjectivity profiles of Dutch causal connectives are very well studied. Dutch differs in its use of causal connectives from English. *Because*, the most frequent English causal connective used in non-basic causal relations, is widely used in both subjective and objective causal relations (Ford 1993, Sweetser 1990). Dutch, on the other hand, tends to use specific causal connectives depending on the subjectivity of the relation. *Want* is the most frequent subjective causal connective; objective relations are most frequently marked by *omdat* (overall, *omdat* is the most generic backward causal connective) (e.g., Sanders & Spooren 2015). When translating relations marked by *because* into Dutch, a choice has to be made between using a subjective or an objective causal connective. Although translators are probably not consciously aware of the difference in subjectivity between the Dutch causal connectives, it can be expected that subjective causal connectives will most often be used to translate subjective *because* relations and that objective *because* relations will be most often translated using an objective causal connective. This was indeed the case in our corpus (see Table 3, $\chi^2(1)=38.85, p<.001$).¹³

Subjective causal relations tend to contain subjective words in their antecedent. We annotated the antecedents of all *because* relations in our corpus to determine the presence of subjective cues ($\kappa=.78$, AC1=.83).¹⁴ The subjective *because* relations indeed often contained subjective words; the objective *because* relations usually did not contain subjective cues (see Table 3, $\chi^2(1)=202.38, p<.001$).¹⁵

¹³ Although Table 3 shows that objective causal connectives tend to be used to translate *because* in objective relations and subjective causal connectives tend to translate *because* in subjective relations, the distribution of subjective and objective connectives over subjective and objective relations is probably not identical to patterns found in non-translated data (see also Cartoni, Zufferey, Meyer & Popescu-Belis 2011, Degand 2004). Translations of *because* from English into Dutch are expected to be biased toward *omdat* (= objective), because it is syntactically equivalent to *because* (both subordinating conjunctions; *want* is a coordinating conjunction) and because it is the most frequent and most general Dutch connective.

¹⁴ At this point, we did not distinguish between different types of subjective cues, but this may be worth exploring in the future. Wei (2018), for instance, shows that the subjective Mandarin Chinese connective *kejian* ‘so’ collocates more often with subjective elements that can be classified as indicating “attitudinal stance” such as *importantly* (Conrad & Biber 2000) than with subjective markers of “epistemic stance”, such as modal verbs.

¹⁵ The category of subjective causal relations in the corpus includes relations with 1st person evaluators in the present tense. Since this is a category of relations that is not prototypically subjective (prototypically, subjective relations have an implicit *subject of consciousness*, see e.g., Pander Maat & Sanders 2001, Pander Maat & Degand 2001, Sanders, Sanders & Sweetser 2009), we also ran all analyses without this group of relations. Since the results remained the same, we did not exclude them from the dataset used in the final analysis.

Being an agreement-type of interaction, the subjective cues often appear in addition to a subjective causal connective, even though both the cue and the connective signal that the consequent is subjective, which makes it highly probable that the relation is subjective. In the Dutch translations in (30) and (31), for instance, the relation is marked by *want*, a subjective connective, even though it is already obvious that the relation is constructed in the speaker's mind from, respectively, *het is erg jammer* 'it is a great pity' or the modal verb *moeten* 'must' in S₁.

- (30) EN *[It is a great pity indeed that Commissioner Barnier has been unable to be present here this morning,] **because** [this is a matter within his brief which is causing great concern not only in Scotland and Wales but in other parts of the Union.]*
{ep-00-03-17}
- NL *[Het is erg jammer dat commissaris Barnier hier vanmorgen niet kon zijn,] **want** [dit is een kwestie uit zijn bevoegdhedenpakket die niet alleen Schotland en Wales, maar ook andere regio's uit de Unie grote zorgen baart.]*
- (31) EN *[We must take the matter up in the Staff Regulations,] **because** [it is an important point.]*
- NL *[We moeten het punt opnemen in het Statuut van de ambtenaren van de Europese Gemeenschappen,] **want** [het is een belangrijk punt.]*

For translation, an agreement type of interaction implies that, unlike in division of labor types of interactions, the segment-internal cue and the connective will not be used as substitutes for each other. More specifically, we do not expect the presence of a cue in the antecedent of the TT relation to influence whether a connective will be used and we should expect to see less variation in cues added or removed from the segments between ST and TT. Table 3 presents an overview of the Dutch translations of all subjective *because* relations in the corpus. As predicted, we found that cues and connectives are indeed not used as substitutes for each other, and that ST and TT relations do not differ in whether or not their segments contain subjective cues, as is also illustrated by the examples in (30) and (31); in other words, all *because* relations with one or multiple subjective elements in their consequents also contained subjective elements in the TT (also in the consequents).

Table 3. Dutch translations of objective and subjective *because* relations with and without subjective elements in the consequent

		objective causal connective ¹⁶	subjective causal connective ¹⁷	temporal / additive / no connective
objective causal relation	+ subj. element in ST+TT	21	13	3
	– subj. element in ST+TT	98	19	24
subjective causal relation	+ subj. element in ST+TT	70	90	33
	– subj. element in ST+TT	6	6	2

The presence of a subjective cue cannot predict whether the TT uses a causal connective ($p=1.00$, Fisher's exact test); subjective relations with a subjective element in their consequent receive a causal connective in the TL as often as subjective relations without a subjective cue. Finally, subjective relations with a subjective cue were translated using a subjective connective as often as subjective relations without a subjective cue ($\chi^2(1)=0.01$, $p=.90$). If the interaction between subjective elements and subjective causal relations had been division of labor, relations with a subjective cue would not be translated using a subjective causal connective.

Subjective causal relations thus show a translation pattern very different from the negative and positive conditional relations discussed in Section 3.2. These differences can be explained on the basis of the different types of interactions between segment-internal cues and connectives found in the respective relations. Both are in turn distinct from the final type of interaction we defined, general collocation, as will be shown in the next section.

5 General collocation

There are many segment-internal features that have been associated with specific types of coherence relations in which the feature and the connective do not signal the same primitive values. Since connectives most prototypically associated with a specific type of coherence relation tend to signal primitive values inferred in those types of relations (e.g., positive causal connectives tend to mark positive causal relations), in this type of interaction the segment-internal elements do not seem to explicitly signal the relation at all. Instead, they seem to function as a cue mainly because they often

¹⁶ *omdat* 'because', *doordat* 'because (of the fact that)', *daar* 'for', *door het feit dat* 'because of the fact that'.

¹⁷ *want* 'because/since', *immers* 'after all', *aangezien* 'considering', *namelijk* 'namely', *gezien het feit dat* 'considering the fact that'.

co-occur with a specific type of relation. In other words, the segment-internal elements and the connectives/relations seem to be *general collocations*.

5.1 General collocation between connectives and segment-internal elements

Asr and Demberg (2015) find that negation elements are associated with, among relation types discussed in Section 4.1, REASON and RESULT relations. Here, the meaning of the segment-internal cue does not overlap with the meaning of the connectives that prototypically mark these types of relations (*because, so*), since REASON and RESULT relations both have a positive value for POLARITY. Rather than division of labor or agreement, the type of interaction between negation elements and the connectives associated with these relations can be more appropriately described as a *general collocation*; the segment-internal cue and the connective often appear together in a relation without being semantically related.¹⁸

A general collocation type of interaction can also be found between causal relations and verb tense. Pit (2003) reports that subjective causal relations often appear in present or future tense, while objective causal relations more often connect segments in past tense. As such, past tense often co-occurs with objective causal connectives, such as German *weil* or French *parce que*; subjective causal connectives, such as Dutch *want* or German *denn*, co-occur more often with present or future tense.

Other examples of general collocations can be found in implicit causality (IC) verbs and verbs of transfer. Both types of verbs have been studied extensively when it comes to their effect on coreference patterns, with NP1 IC verbs (e.g., *apologize, disappoint*) preferring its subject to be the subject of the next clause and NP2 IC verbs (e.g., *admire, fire*) preferring its direct object to be the subject of the next clause (e.g., Au 1986, Garvey & Caramazza 1974, Koorneef & van Berkum 2006, Stewart, Pickering & Sanford 2000); transfer verbs (e.g., *hand, give*) tend to prefer continuations about their goals (usually the indirect object), rather than their sources (Rohde, Kehler & Elman 2006, Stevenson, Crawley & Kleinman 1994). It has, however, been shown that the coreference patterns associated with both IC verbs and verbs of transfer are actually mediated by the type of coherence relation in which they occur. The coreference biases of IC verbs are contingent upon causal relations, either basic or non-basic order positive objective causal relations, i.e., RESULT/CONSEQUENCE or REASON/EXPLANATION (e.g., Koorneef & Sanders 2013, Solstad & Bott 2013); the bias toward continuations about the goal of transfer verbs is mostly found in OCCASION relations, which are a specific type of basic order temporal relations (Rohde et al. 2006) or in RESULT relations (Stevenson

¹⁸ *Collocation* is a term widely used in corpus linguistics to refer to “words that occur (with a frequency that is significantly above chance-level) in a given span around the node word” (Gries & Stefanowitsch 2004: 100), with the words that often co-occur together being referred to as ‘collocates.’

et al. 1994).¹⁹ These types of relations are also the types of relations in which segments containing IC verbs or, respectively, verbs of transfer frequently occur. Implicit causality verbs thus function as a cue for positive objective causal relations, and transfer verbs for basic order temporal relations. Neither type of verb, however, explicitly encodes information about the type of relation that should be inferred between the segment in which it occurs and the upcoming segment, nor do they influence which connective is used to mark the relation. It thus seems likely that both verb types, as well as other segment-internal features in general collocation types of interactions, function as a signal for a specific type of coherence relation by virtue of frequent co-occurrence, i.e., collocation; language users are used to seeing the two together and, as such, encountering a segment-internal feature can help them predict or infer the upcoming coherence relation.

5.2 General collocation in translation

In general collocation types of interactions, there is no overlap in the primitive values signaled by the connective and the meaning expressed by the segment-internal element. As such, they cannot replace each other, and the presence of the one is not expected to make the presence of the other redundant. The translation patterns observed for coherence relations with division of labor and agreement types of interactions between their connectives and segment-internal elements are therefore not expected to be observed in general collocation types of interactions.

One of the general collocation types of interactions identified in Section 5.1 holds between negation elements and causal connectives. In translation, there should be very little variation in the presence of negation in the segments of the relation between the ST and the TT; if there is a negation in the ST, it is expected to also appear in the TT, and, conversely, if there is no negation in the segments of the ST relation, negation is not expected to appear in the translation. In addition, the presence of negation in the TT is not expected to result in an absence of the connective, or in the use of a connective that is less explicit than a causal connective (additive or temporal connective). (32) is an example of an English *because* relation with a negation element, along with its translation into all four languages in the corpus. All four translations include the negation and use a causal connective.

(32) EN [We were not very happy with the early versions from the rapporteur] **because** [we felt the targets were there to almost micro-manage the market in renewable energy sources.]
{ep-00-03-29}

DE [Mit den ersten Versionen des Berichterstatters waren wir nicht so zufrieden,] **da** [wir spürten, die Zielsetzungen waren dazu da, den Markt erneuerbarer Energiequellen

¹⁹ OCCASION: Infer a change of state for a system of entities in S_2 , establishing the initial state for this system from the final state of S_1 (see also Hobbs 1990).

kleinzuhalten.]

ES [Las primeras versiones del ponente no nos gustaban demasiado] **porque** [teníamos la sensación de que el objetivo era gestionar al milímetro el mercado de las fuentes de energía renovables.]

FR [Nous n'étions pas très satisfaits des premières versions du rapport] **parce que** [nous estimions que les objectifs n'existaient presque que pour "microgérer" le marché des sources d'énergies renouvelables.]

NL [We waren niet erg ingenomen met de eerste versies van de rapporteur,] **omdat** [de doelstellingen erop gericht leken de markt van hernieuwbare energiebronnen welhaast tot op het kleinste detail te beheersen.]

As can be seen from the overview of all translations of the *because* relations in the corpus in Table 4, translations never introduce or remove negation elements from the segments of the causal relation. There is also no difference between relations with or without negation when it comes to the use of a causal connective versus a less specific or informative connective option (either an underspecifying connective or no connective, $\chi^2(1)=0.02$, $p=.89$). Since this translation pattern did not differ between languages, we grouped all data together.

Table 4. Explicit vs. implicit translations of *because* relations with vs. without negation into Dutch, German, French, and Spanish

	'because'	temporal / additive / no connective
+ negation in ST+TT	276	15
– negation in ST+TT	1171	69

Table 5 gives an overview of the connectives most frequently used to translate *because* into each language. Unlike for agreement types of interactions, in which the presence of a cue can function as a good predictor for the TT connective, as was shown in Section 4.2, causal relations with a negation element are not more associated with one connective over other candidate connectives, in any language ($p>.05$ for all contrasts). In other words, while there are causal connectives that tend to mark causal relations featuring subjective elements, there seem to be no causal connectives in Dutch, German, French, or Spanish that are preferably used in causal relations featuring negation elements.

Table 5. Most frequent connectives ($n>20$) used to translate *because* relations with vs. without negation into Dutch, German, French, and Spanish

	Dutch		German		French		Spanish	
	conn.	n	conn.	n	conn.	n	conn.	n
+ neg. in	<i>omdat</i>	35	<i>weil</i>	35	<i>car</i>	28	<i>porque</i>	59

SL+TL	<i>want</i>	18	<i>denn</i>	21	<i>parce</i>		<i>ya que</i>	3
	<i>aangezien</i>	2	<i>da</i>	12	<i>que</i>	29		
– neg. in	<i>omdat</i>	154	<i>weil</i>	124	<i>car</i>	147	<i>porque</i>	255
SL+TL	<i>want</i>	55	<i>denn</i>	95	<i>parce</i>		<i>ya que</i>	18
	<i>aangezien</i>	21	<i>da</i>	49	<i>que</i>	109		

Relations involving general collocation types of interactions do not show a specific translation pattern, as was expected on the basis of its characteristics. The way they behave in translation is, however, distinct from relations featuring division of labor or agreement types of interactions. This suggests that there are indeed three different ways in which segment-internal features and connectives can interact.

6 Discussion and conclusion

In this paper, we aimed to develop a systematic way of categorizing segment-internal elements as signals of coherence relations. On the basis of the different ways in which elements inside the segments interact with connectives in the marking of coherence relations, we proposed a three-way distinction between division of labor, agreement, and general collocation. In division of labor types of interactions, the connective and the other signal overlap in the meaning they encode, and the presence of one is likely to make (part of) the other redundant; in agreement types of interactions, the connective and the other signal overlap in the meaning they encode, but they are commonly used in addition to each other. In general collocation types of interactions, there is no overlap in the meaning signaled by the connective and the other signal.

The way in which connectives and segment-internal elements interact in the meaning of coherence relations does not only provide new insights into the question of how coherence relations are marked, but also comments on the questions of how and why elements other than connectives can function as signals of coherence relations. In addition, the categorization proposed in this paper has been shown to make meaningful predictions about the way in which connectives are translated.

In division of labor and agreement types of interactions, it is fairly obvious why segment-internal elements function as cues for coherence relations, since they signal (part of) the same meaning the connective signals or, in case the connective is absent, would signal. In general collocations, the segment-internal element does not explicitly encode which coherence relation should be constructed, but seems to function as a cue because it often co-occurs with a specific type of relation (note that the same mechanism may, but need not be at play in division of labor and agreement types of interactions). This frequent co-occurrence may, however, not be coincidental; plausibly, the driving force behind both this frequent co-occurrence and the fact that a segment-internal element can function as a signal for coherence relations is the expectation of a specific type of coherence relation

that is raised upon encountering a specific segment-internal element. Solstad and Bott (2013), for instance, propose that IC verbs carry an empty “explanatory slot” that has to be filled by information from the discourse. If an explanation for the action expressed by the IC verb has not yet been provided, the explanatory information is expected to follow the IC verb. Such a mechanism would not only explain why IC verbs frequently feature in causal relations, i.e., why speakers often produce this combination, but also why IC verbs can function as signals for causal relations, i.e., why listeners expect and/or look to infer causal relations after an IC verb.

Similar mechanisms can be thought of for the other examples of general collocations discussed in Section 5.1. Sentences involving a negation, for example, often express that something did not happen, which is most relevant if the expectation was that it was going to happen (e.g., Jordan 1988). Why the event did not take place, or why the speaker knows it did not may thus warrant some explanation. Alternatively, an utterance featuring a negation might express that the speaker does not intend to do something or does not like something. Since such messages are usually not what the listener wants to hear, they will often be accompanied by an explanation (e.g., Clayman 2002, Pomerantz & Heritage 2013). Finally, PROBLEM-SOLUTION relations, a specific type of causal relations, involve a negatively evaluated situation, i.e., the problem (e.g., Hoey 1983, Sanders & Noordman 2000, Sanders, Spooren & Noordman 1993). As such, the segment expressing the problem often contains negation elements (Jordan 1984). In sum, there may be a reason for why a segment-internal feature and a specific type of coherence relation frequently co-occur. This underlying mechanism then does not only explain why speakers often produce the two elements together, but also provides an additional explanation as to why – on top of expectations formed on the basis of plain frequency of co-occurrence – segment-internal elements can function as cues for a coherence relation in general collocation types of interactions.

When it comes to the classification of interactions between segment-internal elements and connectives or coherence relations, a potential additional distinction could thus be made between collocations in which there is an underlying cognitive explanation for the frequent co-occurrence of a segment-internal element and a relation type, and collocations in which there is not. Such a distinction, however, would require a lot more research into this phenomenon. In addition, it remains to be determined whether such a distinction would improve the explanatory power of the classification.

Throughout the paper, we discussed many segment-internal elements that have been associated with specific types of coherence relations. The ways in which those observations were made, however, vary from monolingual corpus-based methods (e.g., Asr & Demberg 2015, Pit 2003, Webber 2013), experimental work (implicit causality, verbs-of-transfer), and theoretical explorations (e.g., Carlson 2014, Kehler 1994, Umbach 2005). This yields the question of how the signaling of coherence relations by segment-internal elements is most effectively studied. Corpus-based methods may present the most ecologically valid way of studying signals, but, as already discussed in Section 1.2,

categorization of signals may be largely trivial since there is no obvious one-to-one mapping of signals onto relations. In addition, signals may not be frequent enough to study extensively or locate at all. Corpus-based methods will therefore need to be supplemented with theoretical and experimental work, since these methods allow researchers to zoom in on a single aspect of language. In addition, while theoretical explorations and monolingual corpus studies are powerful tools for formulating hypotheses about language use, experimental methods are more equipped to demonstrate the cognitive plausibility of language models. A clear processing prediction that can be made on the basis of this paper is, for instance, that in the presence of a segment-internal cue, connectives will be less beneficial to readers in division of labor types of interactions (in which case they could even be disruptive) than in agreement or general collocation types of interactions. Using different, complementary approaches can result in a comprehensive and complete overview of a linguistic phenomenon. In this paper, we have demonstrated the usefulness of yet another approach, the use of translation corpora, in studying how coherence relations are signaled. Parallel corpora present a valuable additional tool to research segment-internal signals of coherence relations, especially when it comes to division of labor and agreement types of interactions; it is in these two types of interactions that translation can help make distinctions that in monolingual corpus data would largely be left to the interpretation of the researcher.

In general, relations that contain a segment-internal signal are hypothesized to less often contain a connective than relations that do not contain a segment-internal signal. However, as the current paper argued, not all non-connective signals for coherence relations are created equal. The presence of a segment-internal signal is a much stronger predictor for the absence of the connective in division of labor types of interactions than in agreement or general collocation types of interactions. While the presence of a segment-internal signal may still increase the likelihood of more relations without a connective in agreement and general collocation types of interactions, the mechanism appears much weaker here and more susceptible to other factors influencing the explicit versus implicit marking of coherence relations. Such factors for instance include the segment-internal signal being more strongly associated with another type of coherence relation (e.g., negation is a stronger cue for CHOSEN ALTERNATIVE relations than for causal relations, so a causal relation containing a negation may require a connective to block the expectation of a CHOSEN ALTERNATIVE relation), language users' default expectations about upcoming coherence relations (see Hoek, Zufferey, Evers-Vermeul & Sanders 2017), or the relations' position in the hierarchical discourse structure (see Hoek, Zufferey, Evers-Vermeul & Sanders 2017, Patterson & Kehler 2013, van der Vliet & Redeker 2014). Further examination of how different factors that appear to influence the marking of coherence relations by means of a connective work together seems imperative to fully understanding when language users use connectives to mark coherence relations in a discourse.

References

- Asr, Fatemeh T. & Vera Demberg. 2015. Uniform information density at the level of discourse relations: Negation markers and discourse connective omission. *Proceedings of the 11th International Conference on Computational Semantics (IWCS2015)*, 118–128.
- Au, Terry K. 1986. A verb is worth a thousand words: The causes and consequences of interpersonal events implicit in language. *Journal of Memory and Language* 25(1). 104–122.
- Bazzanella, Carla. 2011. Redundancy, repetition, and intensity in discourse. *Language Sciences* 33(2). 243–254.
- Blakemore, Diane, 1987. *Semantic constraints on relevance*. Oxford: Blackwell.
- Canestrelli, Anneloes R. 2013. *Small words, big effects? Subjective versus objective causal connectives in discourse processing*. Utrecht: University of Utrecht LOT PhD thesis.
http://www.lotpublications.nl/Documents/325_fulltext.pdf
- Carlson, Katy. 2014. Predicting contrasts in sentences with and without focus marking. *Lingua* 150. 78–91.
- Carlson, Lynn, Mary Ellen Okurowski & Daniel Marcu. 2002. *RST Discourse Treebank*. Linguistic Data Consortium, University of Pennsylvania.
- Cartoni, Bruno, Sandrine Zufferey & Thomas Meyer. 2013. Using the Europarl corpus for cross-linguistic research. *Belgian Journal of Linguistics* 27(1). 23–42.
- Cartoni, Bruno, Sandrine Zufferey, Thomas Meyer & Andrei Popescu-Belis. 2011. How comparable are parallel corpora? Measuring the distribution of general vocabulary and connectives. *Proceedings of the 4th workshop on building and using comparable corpora*, 78–86.
- Clayman, Steven E. 2002. Sequence and solidarity. In Shane R. Thye & Edward Lawler (eds.), *Advances in group processes*, vol. 19, 229–253. Bradford: Emerald Group Publishing Limited.
- Conrad, Susan & Douglas Biber. 2000. Adverbial marking of stance in speech and writing. In Geoff Thompson & Susan Hunston (eds.), *Evaluation in text: Authorial stance and the construction of discourse*, 56–74. Oxford: Oxford University Press.
- Das, Debopam & Maite Taboada. 2013. Explicit and implicit coherence relations: A corpus study. *Proceedings of the 2013 annual conference of the Canadian Linguistic Association*.
http://homes.chass.utoronto.ca/~cla-acl/actes2013/Das_and_Taboada-2013.pdf
- Das, Debopam & Maite Taboada. 2018. Signalling of coherence relations in discourse, beyond discourse markers. *Discourse Processes* 55(8). 743–770.
- Das, Debopam, Maite Taboada & Paul McFetridge. 2015. *RST Signalling Corpus LDC2015T10*. Web Download. Philadelphia: Linguistic Data Consortium.

- Degand, Liesbeth. 2004. Contrastive analyses, translation and speaker involvement: The case of *puisque* and *aangezien*. In Michel Achard & Suzanne Kemmer (eds.), *Language, culture and mind*, 251–270. Stanford: CSLI Publications.
- Dyvik, Helge. 1998. A translational basis for semantics. In Stig Johansson & Signe Oksefjell (eds.), *Corpora and cross-linguistic research: Theory, method, and case studies*, 51–86. Amsterdam: Rodopi.
- Evers-Vermeul, Jacqueline, Jet Hoek & Merel C. J. Scholman. 2017. On temporality in discourse annotation: Theoretical and practical considerations. *Dialogue & Discourse* 8(2). 1–20.
- Ford, Cecelia E. 1993. *Grammar in interaction: Adverbial clauses in American English conversations*. Cambridge: Cambridge University Press.
- Frank, Austin & T. Florian Jaeger. 2008. Speaking rationally: Uniform information density as an optimal strategy for language production. *Proceedings of the Cognitive Science Society* 30(30). 939–944.
- Garvey, Catherine & Alfonso Caramazza. 1974. Implicit causality in verbs. *Linguistic Inquiry* 5(3). 459–464.
- Grice, Paul. 1975. Logic and conversation. In Peter Cole & Jerry L. Morgan (eds.), *Syntax and semantics 3: Speech acts*, 41–58. New York: Academic Press.
- Gries, Stefan Th. & Anatol Stefanowitsch. 2004. Extending collocation analysis. *International Journal of Corpus Linguistics* 9(1). 97–129.
- Grisot, Cristina & Joanna Blochowiak. Submitted. *Online processing of temporal discourse relations: Evidence from French using self-paced reading experiments*.
- Halliday, Michael A. K. & Ruqaiya Hasan. 1976. *Cohesion in English*. London & New York: Routledge.
- Halverson, Sandra. 2004. Connectives as a translation problem. In Harald Kittel, Armin Paul Frank, Norbert Greiner, Theo Hermans, Werner Koller, José Lambert & Fritz Paul (eds.), *An international encyclopedia of translation studies*, 562–572. Berlin/New York: Walter de Gruyter.
- Hansen-Schirra, Silvia, Stella Neumann & Erich Steiner. 2007. Cohesive explicitness and explicitation in an English-German translation corpus. *Languages in Contrast* 7(2). 241–265.
- Hansen-Schirra, Silvia, Stella Neumann & Erich Steiner (eds.). 2012. *Cross-linguistic corpora for the study of translations: Insights from the language pair English-German*. Berlin: de Gruyter Mouton.
- Hobbs, Jerry R. 1990. *Literature and cognition*. Stanford: CSLI.
- Hoek, Jet, Jacqueline Evers-Vermeul & Ted J. M. Sanders. submitted. *Using the cognitive approach to coherence relations for discourse annotation*.

- Hoek, Jet, Sandrine Zufferey, Jacqueline Evers-Vermeul & Ted J. M. Sanders. 2017. Cognitive complexity and the linguistic marking of coherence relations: A parallel corpus study. *Journal of Pragmatics* 121. 113–131.
- Hoey, Michael. 1983. *On the surface of discourse*. London: George Allen & Unwin.
- Horn, Laurence R. 1984. Toward a new taxonomy for pragmatic inference. In Deborah Schiffrin (ed.), *Meaning, form, and use in context: Linguistic applications*, 11–42. Washington, D.C.: Georgetown University Press.
- Horn, Laurence R. 1993. Economy and redundancy in a dualistic model of natural language. *SKY 1993: 1993 Yearbook of the Linguistic Association of Finland*. 33–72.
- Jordan, Michael P. 1984. *Rhetoric of everyday English texts*. London: George Allen & Unwin.
- Jordan, Michael P. 1988. The power of negation in English: Text, context and relevance. *Journal of Pragmatics* 29(6). 705–752.
- Kehler, Andrew. 1994. Temporal relations: Reference or discourse coherence? *Proceedings of the 32nd Annual Meeting on Association for Computational Linguistics*, 319–321.
- Knott, Alistair & Robert Dale. 1994. Using linguistic phenomena to motivate a set of coherence relations. *Discourse Processes* 18(1). 35–62.
- Knott, Alistair & Ted J. M. Sanders. 1998. The classification of coherence relations and their linguistic markers: An exploration of two languages. *Journal of Pragmatics* 30(2). 135–175.
- Koehn, Phillip. 2005. Europarl: A parallel corpus for statistical machine translation. *Tenth Machine Translation Summit (MT Summit X)*, <http://homepages.inf.ed.ac.uk/pkoehn/publications/europarl-mtsummit05.pdf> (8 April, 2014).
- Koornneef, Arnout W. & Jos J. A. van Berkum. 2006. On the use of verb-based implicit causality in sentence comprehension: Evidence from eye tracking. *Journal of Memory and Language* 54(4). 445–465.
- Koornneef, Arnout W. & Ted J. M. Sanders. 2013. Establishing coherence relations in discourse: The influence of implicit causality and connectives on pronoun resolution. *Language and Cognitive Processes* 28(8). 1169–1206.
- Lascarides, Alex & Nicholas Asher. 1993. Temporal interpretation, discourse relations and commonsense entailment. *Linguistics and Philosophy* 16(5). 437–493.
- Levshina, Natalia & Liesbeth Degand 2017. Just *because*: In search of objective criteria of subjectivity expressed by causal connectives. *Dialogue & Discourse* 8(1). 132–150.
- Levy, Roger & T. Florian Jaeger. 2007. Speakers optimize information density through syntactic reduction. In Bernhard Schölkopf, John Platt & Thomas Hoffman (eds.), *Advances in neural information processing systems (NIPS)*, vol. 19, 849–856. Cambridge, MA: MIT Press.
- Li, Fang, Ted J. M. Sanders & Jacqueline Evers-Vermeul. 2016. On the subjectivity of Mandarin reason connectives: Robust profiles or genre-sensitivity? In Ninke M. Stukker, Wilbert P. M. S.

- Spooren & Gerard J. Steen (eds.), *Genre in language, discourse and cognition*, 15–49. Amsterdam: de Gruyter Mouton.
- Mann, William C. & Sandra A. Thompson. 1988. Rhetorical Structure Theory: Toward a functional theory of text organization. *Text* 8(3). 243–281.
- Melamed, I. Dan. 2001. *Empirical methods for exploiting parallel texts*. Cambridge, MA: MIT Press.
- Noël, Dirk. 2003. Translations as evidence for semantics: An illustration. *Linguistics* 41(4). 757–785.
- Pander Maat, Henk L. W. & Liesbeth Degand. 2001. Scaling causal relations and connectives in terms of speaker involvement. *Cognitive Linguistics* 12(3). 211–246.
- Pander Maat, Henk L. W. & Ted J. M. Sanders. 2001. Subjectivity in causal connectives: An empirical study of language in use. *Cognitive Linguistics* 12(2). 247–274.
- Patterson, Gary & Andrew Kehler. 2013. Predicting the presence of discourse connectives. *Proceedings of the 2013 Conference on Empirical Methods in Natural Language Processing (EMNLP2013)*. 914–923.
- Prasad, Rashmi, Eleni Miltsakaki, Nikhil Dinesh, Alan Lee, Aravind K. Joshi, Livio Robaldo & Bonnie L. Webber. 2008. The Penn Discourse TreeBank 2.0. *Proceedings of the 6th International Conference on Language Resources and Evaluation (LREC 2008)*. 2961–2968.
- Pit, Mirna. 2003. *How to express yourself with a causal connective: Subjectivity and causal connectives in Dutch, German, and French*. Amsterdam: Rodopi.
- PDTB Research Group. 2007. *The Penn Discourse Treebank 2.0 annotation manual*. IRCS technical report. http://repository.upenn.edu/cgi/viewcontent.cgi?article=1203&context=ircs_reports
- Pomerantz, Anita & John C. Heritage. 2013. Preference. In Jack Sidnell & Tanya Stivers (eds.), *The handbook of conversation analysis*, 210–228. Chichester: John Wiley & Sons.
- R Core Team, 2016. *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. <https://www.R-project.org/>
- Renkema, Jan. 2009. *The texture of discourse: Towards an outline of connectivity theory*. Amsterdam: John Benjamins.
- Reese, Brian, Julie Hunter, Nicholas Asher, Pascal Denis & Jason Baldridge. 2007. *Reference manual for the analysis of rhetorical structure*. Unpublished manuscript. University of Texas at Austin, Austin, TX. http://timeml.org/jamesp/annotation_manual.pdf
- Rohde, Hannah, Andrew Kehler & Jeffrey L. Elman. 2006. Event structure and discourse coherence biases in pronoun interpretation. *Proceedings of the 28th Annual Conference of the Cognitive Science Society*. 697–702.
- Sanders, Ted J. M. & Leo G. M. Noordman. 2000. The role of coherence relations and their linguistic markers in text processing. *Discourse Processes* 29(1). 37–60.
- Sanders, Ted J. M., José Sanders, Eve E. Sweetser. 2009. Causality, cognition and communication: A mental space analysis of subjectivity in causal connectives. In Ted J. M. Sanders & Eve E.

- Sweetser (eds.), *Causal categories in discourse and cognition*, 19–60. Berlin: Mouton de Gruyter.
- Sanders, Ted J. M. & Wilbert P. M. S. Spooren. 2007. Discourse and text structure. In Dirk Geeraerts & Hubert Cuyckens (eds.), *The Oxford handbook of cognitive linguistics*, 916–941. Oxford: Oxford University Press.
- Sanders, Ted J. M. & Wilbert P. M. S. Spooren. 2015. Causality and subjectivity in discourse: The meaning and use of causal connectives in spontaneous conversation, chat interactions and written text. *Linguistics* 53(1). 53–92.
- Sanders, Ted J. M., Wilbert P. M. S. Spooren & Leo G. M. Noordman. 1992. Toward a taxonomy of coherence relations. *Discourse Processes* 15(1). 1–35.
- Sanders, Ted J. M., Wilbert P. M. S. Spooren & Leo G. M. Noordman. 1993. Coherence relations in a cognitive theory of discourse representation. *Cognitive Linguistics* 4(2). 93–133.
- Solstad, Torgrim & Oliver Bott. 2013. Towards a formal theory of explanatory biases in discourse. *Proceedings of the 19th Amsterdam Colloquium*, 203–210.
- Sperber, Dan & Deirdre Wilson. 1985. Loose talk. *Proceedings of the Aristotelian Society LXXXVI*, 540–549.
- Spooren, Wilbert P. M. S. 1997. The processing of underspecified coherence relations. *Discourse Processes* 24(1). 149–168.
- Stevenson, Rosemary J., Rosalind A. Crawley & David Kleinman. 1994. Thematic roles, focus and the representation of events. *Language and Cognitive Processes* 9(4). 519–548.
- Stewart, Andrew J., Martin J. Pickering & Anthony J. Sanford. 2000. The time course of the influence of implicit causality information: Focusing versus integration accounts. *Journal of Memory and Language* 42(3). 423–443.
- Stukker, Ninke M., Ted J. M. Sanders & Arie Verhagen. 2008. Causality in verbs and in discourse connectives: Converging evidence of cross-level parallels in Dutch linguistic categorization. *Journal of Pragmatics* 40(7). 1296–1322.
- Sweetser, Eve E. 1990. *From etymology to pragmatics: The mind-body metaphor in semantic structure and semantic change*. Cambridge: Cambridge University Press.
- Taboada, Maite & Debopam Das. 2013. Annotation upon annotation: Adding signalling information to a corpus of discourse relations. *Dialogue and Discourse* 4(2). 249–281.
- Tanskanen, Sanna-Kaisa. 2006. *Collaborating towards coherence: Lexical cohesion in English discourse*. Amsterdam: John Benjamins.
- Teubert, Wolfgang. 1999. Corpus linguistics: A partisan view. *TELRI Newsletter* 8. 4–19.
- Umbach, Carla. 2005. Contrast and information structure: A focus-based analysis of *but*. *Linguistics* 43(1). 207–232.

- Vliet, Nynke van der & Gisela Redeker. 2014. Explicit and implicit coherence relations in Dutch texts. In Helmut Gruber & Gisela Redeker (eds.), *The pragmatics of discourse coherence: Theories and applications*, 23–52. Amsterdam/Philadelphia: John Benjamins.
- Webber, Bonnie L. 2013. What excludes an alternative in coherence relations? *Proceedings of the 10th International Workshop on Computational Semantics (IWCS2013)*, 276–287.
- Wei, Yipu. 2018 *Causal connectives and perspective markers in Chinese: The encoding and processing of subjectivity in discourse*. Utrecht: University of Utrecht LOT PhD thesis.
https://www.lotpublications.nl/Documents/482_fulltext.pdf
- Wilson, Deirdre & Dan Sperber. 1993. Linguistic form and relevance. *Lingua* 90(1–2). 1–25.
- Wilson, Deirdre & Dan Sperber. 2005. Relevance Theory. In Laurence R. Horn & Gergory Ward (eds.), *The handbook of pragmatics*, 607–632. New York: Wiley.
- Zufferey, Sandrine & Bruno Cartoni. 2014. A multifactorial analysis of explicitation in translation. *Target* 26(3). 23–42.

Authors' addresses

Jet Hoek
 jet.hoek@ed.ac.uk
 Dugald Stewart Building
 3 Charles Street
 Edinburgh, EH8 9AD
 United Kingdom

Sandrine Zufferey
 sandrine.zufferey@rom.unibe.ch
 Institut de langue et de littérature françaises
 Länggassstrasse 49
 3012 Berne
 Switzerland

Jacqueline Evers-Vermeul
 j.evers@uu.nl
 Trans 10
 3512 JK Utrecht
 The Netherlands

Ted J.M. Sanders
t.j.m.sanders@uu.nl
Trans 10
3512 JK Utrecht
The Netherlands

Biographical notes

Jet Hoek (PhD Utrecht University 2018) is a post-doctoral researcher at the University of Edinburgh. She uses a range of methodologies (both on-line and off-line) to research discourse-level phenomena, focusing mainly on coherence relations and referring expressions.

Sandrine Zufferey (born 1978, PhD University of Geneva 2007) is full professor of French linguistics at the University of Bern in Switzerland. Her current research focuses on the way native and non-native speakers use, understand and process discourse connectives, as well as on the way discourse relations are encoded across languages and in translations.

Jacqueline Evers-Vermeul, assistant professor at the Department of Languages, Literature and Communication at Utrecht University, studies the representation, production, comprehension and teaching of text-structural aspects of discourse. She uses a variety of methods to study adult connective use (synchronic and diachronic corpus-based analyses), students' understanding of school texts and examination questions (on- and off-line reading experiments among children and teenagers), and children's connective acquisition (visual-world experiments among 2- and 3-year-olds, growth-curve analyses of dense longitudinal corpora).

Ted J.M. Sanders (Ph.D. 1992) is a full professor of Discourse Studies at the Department of Languages, Literature and Communication and the Utrecht Institute of Linguistics of Utrecht University, The Netherlands. His research concentrates on the (cross-)linguistic realization of coherence in discourse, as well as on the cognition of discourse processing and representation. Furthermore, he has keen interest in text comprehensibility and communication. With co-authors he published in many international journals and books. He co-edited special issues of journals and two book volumes. He is currently vice-dean of the Faculty of Humanities.

Date accepted: Dec. 1st, 2018