

# Says who? On the treatment of speech attributions in discourse structure

Gisela Redeker & Markus Egg, University of Groningen<sup>1</sup>

**Abstract:** In this paper, we present what we think is an elegant solution to some problems in the discourse-structural modelling of speech attribution. Using mostly examples from the Wall Street Journal Corpus, we show that the approach proposed by Carlson and Marcu (2001) leads to irresolvable dilemmas that can be avoided with a suitable treatment of attribution in an underspecified representation of discourse structure.

Most approaches to discourse structure assume that textual coherence can be modelled as trees. In particular, it has been shown that coherent discourse follows the so-called *right-frontier constraint (RFC)*, which essentially ascertains a hierarchical structure without crossed dependencies. We will discuss putative counterexamples to these two assumptions, most of which involve reported speech as in (1) (cited in Wolf and Gibson 2005):

- (1) “Sure I’ll be polite,” promised one BMW driver who gave his name only as Rudolph. “As long as the trucks and the timid stay out of the left lane.”

In (1) the second part of the quote should be linked to the first part (and not the whole first sentence) by a *condition* relation. If we were to analyse the parenthetical speech reporting clause (“promised one BMW driver ...”) as the nucleus of its host clause (i.e., the quote), the RFC would prevent linkage between the two parts of the quote. If the attribution is analysed as a satellite of the quote, as in Carlson and Marcu (2001), Wolf and Gibson argue, it should be a satellite to *both* parts of the quote, thus violating treeness.

In this paper, we will explore the problems arising from this type of construction and propose a treatment of speech report attributions that we will argue allows us to preserve both, treeness and the RFC in building discourse structures.

## 1 The (non-)treatment of speech attribution in classic Rhetorical Structure Theory (RST)

In ‘classic’ Rhetorical Structure Theory (RST; Mann and Thompson 1988), the problems in accommodating speech report attributions do not arise, because classic RST does not separate complements of verbs and parenthetical speech reporting clauses from their host clause.

Leaving speech attribution implicit is in line with the general ‘philosophy’ of RST, which aims to represent not all possible links, but the most plausible structure licensed by

---

<sup>1</sup> Markus Egg is now at Humboldt University, Berlin. This manuscript is a slightly updated (2009) version of our paper in the *Proceedings of the Workshop on Constraints in Discourse, Maynooth, Ireland 2006* (<http://www.constraints-in-discourse.org/cid06/>).

the interpretation of the segments. A case in point is the conflation of semantic (*ideational, informational, or subject-matter*) and pragmatic (*interpersonal and textual, intentional, or presentational*) structure in RST: For any two segments, the analyst has to choose a single relation, although most pragmatic relations rely on a simultaneously realised semantic relation (Redeker 2000: 248f). Unlike other authors (e.g. Moore and Pollack 1992), we consider this an adequate reflection of the fact that texts usually consist of informational (narrative or expository) stretches and varying amounts of intentional or reader-oriented material. The analyst (and supposedly also the reader) selects the most plausible relation in the context at hand, ignoring an underlying semantic relation where a pragmatic one is more salient.

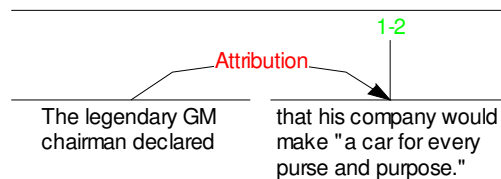
One motivation for not analysing speech report attributions is the notorious ambiguity of unmarked subsequent sentences that are often clearly informed by the same source (e.g., stated in the same interview or press conference), but reported factually and without quotation marks. Their status may be genuinely undecidable. We will return to this issue below.

Still, relegating the treatment of quotation entirely to another type of analysis (e.g. *Engagement*; see White 2003) seems unsatisfactory for newspaper corpora, where explicit attribution of information to sources is a very salient genre-specific feature with arguable relevance for discourse structure.

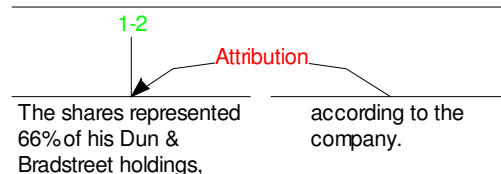
## 2 Carlson and Marcu's (2001) treatment of speech attribution

A systematic treatment of speech attribution has been proposed in the annotation scheme of Carlson and Marcu (2001). In this scheme, the reporting clause of a direct quote and the matrix clause of an indirect speech report are attached to the segment(s) presenting the quoted or reported speech as the satellite of a relation of *attribution*, as illustrated in examples (2a) and (2b) from Marcu's Wall Street Journal Corpus:

(2) a. wsj\_1377



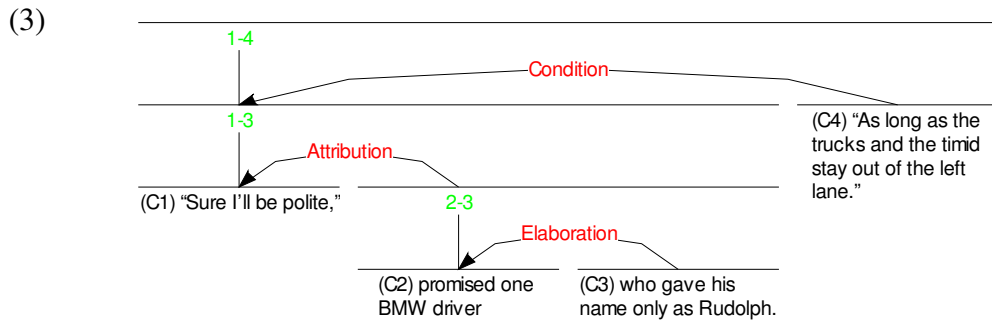
b. wsj\_1157



We identify several problems with Carlson and Marcu's proposal. First of all, the attribution relation is defined not only for speech verbs, but also for 'cognitive' predicates. The motivation for this is evident: *verba cogitandi* such as *to believe, to expect, etc.* are often used (in the form of grammatical metaphor) in speech report

clauses. But the attribute relation is applied indiscriminately for speech verb uses *and* for mental verb uses. Moreover, all these attributed segments are assigned nuclear status and are thus not dominated by a node that might block unwarranted inferences from these opaque contexts.

Another configurational problem concerns the quite frequent case of attribution satellites that interrupt a quotation, as in example (1) above. A Marcu-style analysis of this example is given in (3).



By assigning satellite status to the reporting clause, this analysis allows the two parts of the quotation,  $C_1$  and  $C_4$ , to be linked with the condition relation without violating the RFC (under which the latest segment and all nodes dominating it remain accessible). At the same time, this analysis creates the problem noted by Wolf and Gibson (2005) that the reporting clause cannot be linked to  $C_4$  without violating the treeness constraint that bars multiple parent nodes.<sup>2</sup> Without such a link, this representation cannot differentiate between cases where  $C_4$  is within the scope of the attributing clause and those where it is not, as  $C_4$  is related to  $C_1$ - $C_3$  and not to  $C_1$  alone. This could only be avoided if the analysis assigned nuclear status to the medial reporting clause, thus allowing both parts of the quotation as satellites. Such an analysis, however, would preclude any direct relation between  $C_1$  and  $C_4$ , and would thus not yield a plausible analysis of this example. (We will return to this point in Section 3.)

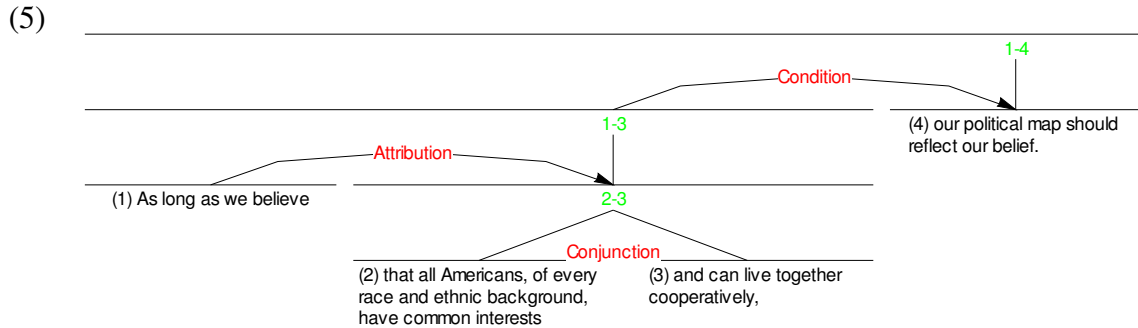
Another configurational problem with Carlson and Marcu's rule to assign satellite status to the reporting clause arises for examples like (4), where the cognitive predicate in  $C_1$  and not the attributed material (here  $C_2$ - $C_3$ ) provides the link for the next segment (here  $C_4$ ).

- (4) As long as we believe ( $C_1$ ) that all Americans, of every race and ethnic background, have common interests ( $C_2$ ) and can live together cooperatively ( $C_3$ ), our political map should reflect our belief ( $C_4$ ). (wsj\_1137)

Intuitively,  $C_4$  states a consequence of the fact that a group including the speaker believes  $C_2$ - $C_3$  (rather than stating a consequence of  $C_2$ - $C_3$  themselves). This is supported by the lexical repetition (*believe* – *belief*). However, in the Marcu-style analysis (5) of this example, this intuition cannot be modelled, because no discourse relation can link  $C_1$  (the

<sup>2</sup> This problem arises only for quotes consisting of more than one atomic discourse segment. If an atomic segment is interrupted by a parenthetical, Carlson and Marcu use a quasi-relation SAME-UNIT, whose only purpose is to express the fact that the two parts of the interrupted segment belong together.

belief-attributing clause) to  $C_4$  (the consequence), as  $C_1$  is a satellite of  $C_2$ - $C_3$  and is thus no longer at the right frontier. Linking  $C_1$  to  $C_4$  would be possible only if  $C_1$  was the nucleus, and  $C_2$ - $C_3$ , its satellite.



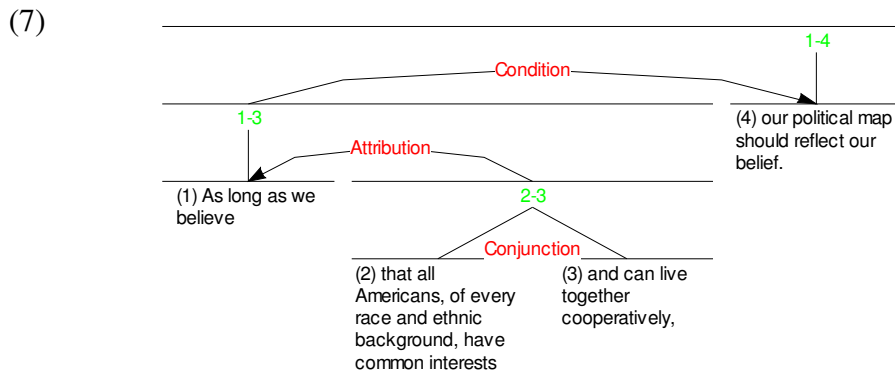
This problem is not restricted to belief attributions, but also occurs with quotation and reported speech. In (6), for example, “tried to put the blame” is expanded on by “saying” and “wrote” and only indirectly by what Mr. Trump said and wrote.

(6) Yesterday, Mr. Trump tried to put the blame for the collapse of the UAL deal on Congress, *saying* it was rushing through a bill to protect AMR executives. “I believe that the perception that legislation in this area may be hastily approved contributed to the collapse of the UAL transaction, and the resulting disruption in the financial markets experienced this past Friday,” Mr. Trump *wrote* members of Congress. (from *wsj\_file5*; italics added)

Here again, then, it would seem advantageous to have an attribution relation that assigns nuclear status to the reporting clause.

### 3 A proposal

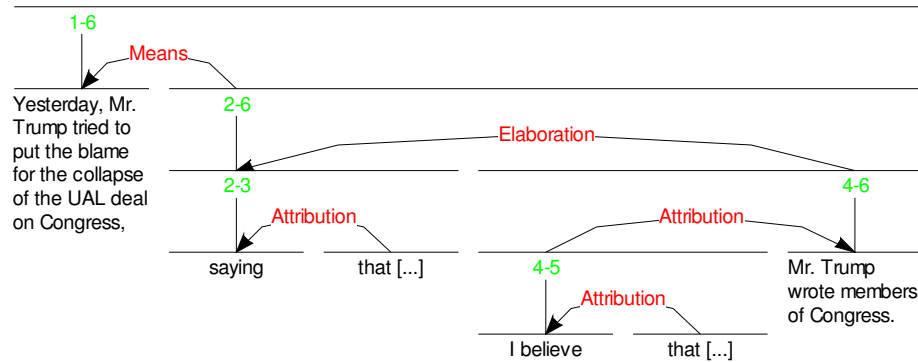
In the light of the above discussion, we propose treating speech reporting and belief attributing clauses as the nuclei of attribution relations. The structure for example (4) would then be (7), which clearly highlights the link between  $C_1$  and  $C_4$  (as argued for in Section 2).



Similarly, example (6) could be assigned the global structure shown in (8). For cognitive

predications, modal and negated context, representing the complements as satellites would more intuitively reflect their opacity, and clearly demarcate the scope of the attribution.

(8)



That still leaves us with the problem of parenthetical attributions in quote-medial position, as illustrated with example (1). Here we propose to adapt the treatment of parenthetical segments common in classic RST (suggested, e.g., by O’Donnell 1997): We move message-internal speech attributions to the *beginning* of the message, which is their satellite in an attribution relation. This solves the two problems: They are moved from their position in order not to interrupt the constituent they are embedded in. Second, the range of the message can be specified exactly in terms of their satellite, which is particularly important for ‘cognitive predicates’ such as *believe* that introduce a modal context.

Consideration of examples like (9) and (10) below suggests that we should even go one step further and move not only sentence-medial, but also sentence-final attributions to the beginning of the sentence they occur in.

- (9) “When Sears has a sale at a special price,” the woman in the ad declares, “it’s something you don’t want to miss.” (wsj\_1105)
- (10) “We’ve got a lot to do,” he acknowledged. “We’ve got to move quickly.” (wsj\_0616)

Note that this changes the analysis of example (6) given in (8): the segments C<sub>4</sub>-C<sub>5</sub> and C<sub>6</sub> will now occur in reversed order, allowing for the possibility of the quotation to be continued in a following sentence.

### 3.1 Underspecified representation

The analysis we propose can be integrated straightforwardly into the syntax-discourse interface proposed in Egg and Redeker (2008) for the automatic derivation of information on discourse structure from syntactic structure. In this interface, underspecified descriptions of discourse structure (so-called *constraints*) are updated by integrating new material at a distinguished position (the *immediately preceding discourse segment*) and then determining the new immediately preceding discourse segment. For reporting clauses in medial position that are preceded by part of the attributed material in a segment

C<sub>1</sub>, the attribution relation is introduced in such a way that it links these segments, but that C<sub>1</sub> retains its status as the immediately preceding discourse segment. In that way, any further incoming material can be linked directly to C<sub>1</sub> without the reporting clause C<sub>2</sub> intervening.

This strategy conveniently allows a deliberate underspecification of the discourse structure for cases like (11) below, where the boundary of the message is not clear: Does C<sub>3</sub>-C<sub>4</sub> indicate the reason for the fact that the market makers make a statement (i.e., C<sub>3</sub>-C<sub>4</sub> is a satellite of C<sub>1</sub>-C<sub>2</sub>), or is it part of this statement, and thus a satellite of C<sub>2</sub> only? Instead of simply choosing one of the possibilities at random, there should be a way of expressing that the discourse structure of (11) is ambiguous.

(11) The market makers say (C<sub>1</sub>) they aren't comfortable carrying big positions in stocks (C<sub>2</sub>) because they realize (C<sub>3</sub>) prices can tumble quickly (C<sub>4</sub>). (wsj\_1142)

The resulting structure deliberately leaves unspecified whether C<sub>3</sub> etc. fall within the scope of C<sub>1</sub> and is thus compatible with exactly two discourse structures, one for each interpretation of (11). If the attributive fragment receives wide scope, C<sub>3</sub>-C<sub>4</sub> is a part of the message; otherwise, C<sub>3</sub>-C<sub>4</sub> relates to a discourse segment C<sub>1</sub>-C<sub>2</sub> and gives a reason why the statement is made.

#### 4 What next?

After exploring the empirical constellations of speech attributions with examples from the WSJ corpus and showing that they can be adequately modelled in an underspecified discourse structure representation, we can now move on to test the feasibility of our proposal with a wider range of examples, with different genres, and in other languages. In this concluding paragraph, we will sketch some of the issues we expect to encounter.

##### 4.1 Extended quotation with multiple attributions

Moving all attributions to the beginning of the sentence they occur in might create problems in multi-sentential quotes with multiple attribution phrases. In example (6), for instance, moving C<sub>6</sub> in front of C<sub>4</sub>-C<sub>5</sub> entails that C<sub>4</sub>-C<sub>5</sub> cannot be directly related to previous parts of the quotation (which in this case would not be plausible).

We assume that quote continuations with a sentence-initial attributional phrase should be unproblematic, as the writer clearly presents the added material as a separate (piece of the) quote. Potentially problematic cases would thus be continuations with a medial or final attribution. We identified 21 relevant examples in the 'Test' part of Marcu's WSJ corpus (comprising 38 texts with a total of 2,347 atomic segments). It appears that they can all be plausibly analysed by relating the combined attribution + quote segments instead of just the quoted material.

The most common and, we feel, most obviously unproblematic structure (occurring in ten cases) is a combination of an indirect speech report followed by a direct quote (see example (12)).<sup>3</sup> The initial report gives a general or abstract introduction to or version of what is then detailed, illustrated or evaluated in a direct quote.

---

<sup>3</sup> In examples (12) through (15), the critical attribution phrases (which in our proposal would be moved to the front of their sentence) are italicised.

- (12) Sara Albert, a 34-year-old Dallas law student, says she’s generally skittish about the stock market and the takeover activity that seems to fuel it. “I have this feeling that it’s built on sand,” *she says*, that the market rises “but there’s no foundation to it.”

In eight other cases, the quotation parts are rather loosely related pieces where relations like JOINT or ELABORATION could be applied at the level of the quoted utterances or at the level of the quoting or reporting sentences.

There are only three cases where a stronger relation is discernible, as the continuation begins with the connective *but*. In all three cases, however, we feel that the connective signals a contrast between the reported speech acts (‘A says X, but then A also says Y’) rather than a contrast between the parts of the quotation (‘A says: X but Y’). This intuition is based not only on the presence of the second attributing phrase, but also on the presence of intervening information (a sentence-final attribution after the first part of the quote in examples (13) and (14), and an embedded quotation of a difference source in (15)). Here are all three cases:

- (13) Savings and loans reject blacks for mortgage loans twice as often as they reject whites, the Office of Thrift Supervision said. But that doesn’t necessarily mean thrifts are discriminating against blacks, *the agency said*. The office, an arm of the Treasury, said it doesn’t have data on the financial position of applicants and thus can’t determine why blacks are rejected more often.
- (14) “The opposition can be the most hurt because everyone already figures the LDP is that kind of beast,” says Shigezo Hayasaka, former aide to LDP kingmaker Kakuei Tanaka and now an independent analyst. But, *he adds*, “We can’t tell where it will go at all because we’re still in the middle of it.”
- (15) “The market started with several strikes against it,” said Peter DaPuzzo, head of retail equity trading at Shearson Lehman Hutton, referring to news that the labor-management buy-out of UAL Corp. continued to unravel, and reports that the junk-bond market is disintegrating. But the computer-guided selling in response to those developments dealt a serious blow to the over-the-counter market, *Mr. DaPuzzo said*.

While this limited exploration suggests that our proposal can probably handle extended quotes, more systematic screening of examples, especially from other genres, is needed to be sure this finding can be generalized.

#### 4.2 Cognitive predicates

In Carlson and Marcu (2001), ATTRIBUTION covers speech reports and cognitive predicates (*verba cogitandi* and *verba sentiendi*) alike, as long as the source is specified and the complement is finite. As we move from newspaper texts to other genres, however, we will encounter 1<sup>st</sup>- and 2<sup>nd</sup>-person cognitive predicates, which in the vast majority of cases function not as attributions, but rather as more or less formulaic indicators of the speaker’s epistemic, evidential or evaluative stance. The examples in

(16) are from Thompson (2002), who examined 425 finite indicative complements from a corpus of conversations.

- (16) I remember I was talking to him regularly for a time. (epistemic)  
Let's find out if it works. (evidential)  
I don't give a shit what she thinks. (evaluative)

In such cases, then, what looks like and is generally analysed as complements should probably not be considered as separate atomic segments.

What also will need to be investigated further are cases (also reported in Thompson 2002) where 3<sup>rd</sup>-person predicates seem to function in the same way. We expect this to be a characteristic of informal conversation which will not usually occur in the written text genres that are the main focus of our approach.

### 4.3 To conclude

These explorations suggest that our proposal provides a viable treatment of speech reports in written monologic texts. Whether it also yields adequate representations of belief statements may be a more difficult question. Ultimately, an account of the discourse-structural properties of attribution will contribute to the ongoing debate on the segmentation and configuration of discourse.

### References

- Carlson, L. and D. Marcu (2001). *Discourse tagging reference manual*. (Available: <http://www.isi.edu/~marcu/discourse/tagging-ref-manual.pdf>).
- Egg, M. and G. Redeker (2008). Underspecified discourse representation. In A. Benz and P. Kühnlein (eds.), *Constraints in Discourse* (pp. 117-138). Amsterdam: Benjamins.
- Mann, W. and S. Thompson (1988). Rhetorical Structure Theory: Towards a functional theory of text organization. *Text*, 8, 243-281.
- Marcu, D. (2000). *The theory and practice of discourse parsing and summarization*. Cambridge, MA: MIT Press.
- Moore, J. and M. Pollack (1992). A problem for RST: the need for a multi-level discourse analysis. *Computational Linguistics*, 18, 537-544.
- O'Donnell, M. (1997). RST-Tool: An RST Analysis Tool. *Proceedings of the 6<sup>th</sup> European Workshop on Natural Language Generation 1997*. University of Duisburg. (Available: <http://www.wagsoft.com/Papers/EWNLG-RSTTool.ps>).
- Redeker, G. (2000). Coherence and structure in text and discourse. In W. Black and H. Bunt (eds.), *Abduction, Belief and Context in Dialogue*. *Studies in Computational Pragmatics* (pp. 233-263). Amsterdam: Benjamins.
- Thompson, S.A. (2002). "Object complements" and conversation: towards a realistic account. *Studies in Language*, 26(1), 125-164.
- White, P.R.R. (2003). Beyond modality and hedging: A dialogic view of the language of intersubjective stance. *Text*, 23, 259-284.
- Wolf, F. and E. Gibson (2005). Representing discourse coherence: a corpus-based study. *Computational Linguistics*, 31, 249-287.