Can text structure be incompatible with rhetorical structure?

Nadjet Bouayad-Agha, Richard Power and Donia Scott
Information Technology Research Institute
University of Brighton
Lewes Road
Brighton BN2 4GJ, UK
firstname.lastname@itri.bton.ac.uk

Abstract

Scott and Souza (1990) have posed the problem of how a rhetorical structure (in which propositions are linked by rhetorical relations, but not yet arranged in a linear order) can be realized by a text structure (in which propositions are ordered and linked up by appropriate discourse connectives). Almost all work on this problem assumes, implicitly or explicitly, that this mapping is governed by a constraint on compatibility of structure. We show how this constraint can be stated precisely, and present some counterexamples which seem acceptable even though they violate compatibility. The examples are based on a phenomenon we call *extraposition*, in which complex embedded constituents of a rhetorical structure are extracted and realized separately.

1 Introduction

Text planning (or more broadly, document planning) can be divided into two stages. In the first stage, material is selected, perhaps from a knowledge base, and organized rhetorically. In the second stage, the rhetorical structure is realized by a text structure (or document structure), through which the material is distributed among sentences, paragraphs, vertical lists, and perhaps even diagrams. The RAGS (1999) proposal for a standard NLG architecture distinguishes the outputs of these two phases by the data types `RhetRep` (rhetorical representation) and `DocRep` (document representation).

We focus in this paper on the second stage of text planning — the passage from `RhetRep` to `DocRep`. NLG researchers have addressed this issue in various ways, but everyone assumes some kind of structural compatibility between rhetorical structure and text structure. The most popular discourse framework in NLG is RST (Mann and Thompson, 1988), which makes the crucial distinction between nucleus, which is the most important part of a message, and satellite, which is the peripheral part of the message. Scott and Souza (1990) provide guidelines for the realisation of RST trees into a coherent text. One of them is to avoid dangling sentences, that is, to avoid putting "information that is only weakly relevant to the message" in a separate sentence because it will feel as if it has been introduced as an afterthought or as introducing a new topic which is then abruptly abandoned, disrupting the comprehensibility of the text. As an example, the authors provide the attributive satellite of an elaboration relation.

Marcu (1996), in order to build a valid text plan from a set of rhetorical assertions, uses the "nuclearity principle", that is the observation in Mann and Thompson's framework that "whenever two large text spans are connected through a rhetorical relation, that rhetorical relation holds also between the most important parts of the constituent spans". Therefore, the resulting text plans are valid in the sense that they are isomorphic with one of the rhetorical structures that can be built from the rhetorical assertions using this nuclearity principle.

Our aim in this paper is to formulate more precisely a notion of structural compatibility which is necessary in order to describe the realisation of a `RhetRep` into various `DocReps`, and then to discuss some examples (mostly taken from the domain of patient information leaflets) of apparently acceptable texts in which this notion of compatibility is violated. To discuss this issue clearly, an assumption must be made about the kinds of information represented by rhetorical and text structure; we outline in section 2 the common assumption that these representations are trees, labelled respectively with rhetorical and textual categories, the rhetorical structure being unordered and the text structure ordered. Section 3 then defines a notion of structural compatibility that is weaker than isomorphism; section 4 shows that we can find plausible counterexamples even to this weaker formulation, and discusses why these passages occur. Section 5 discusses some implications for NLG, and finally, section 6 raises further important issues.

2 Rhetorical structure and text structure

To distinguish clearly between `RhetRep` and `DocRep`, we need to define the kinds of information that should be included in the two representations. Bateman and Rondhius (1997) compare several approaches to rhetorical representation, citing in particular RST (Mann and Thompson, 1988) and Segmented Discourse Representation Theory (Asher, 1993). These approaches share the idea that rhetorical representations are composed of propositions linked by rhetorical relations; SDRT includes as well the logical apparatus of DRT, thus covering notions like necessity and logical scope which are missing from RST. For the most part, NLG applications have used the RST framework, adapted in various ways; the most common representation, proposed also as the RAGS standard, is that of a tree in which terminal nodes represent elementary propositions, while non-terminal nodes represent rhetorical relationships. This representation, proposed for example by Scott and Souza (1990), is illustrated by figure 1, which might be realized by the following passage:

(1) Elixir occasionally provokes a mild allergic reaction^B, because it contains gestodene^C. However, Elixir has no serious side-effects^A.

Assuming an RST-based framework, an important issue is whether the rhetorical representation should already imply a linear order. Most researchers have followed Scott and Souza in assuming that linear order should be left unspecified; it is during the transition to the document representation that the material is distributed among linguistic units (or perhaps diagrams, in a multimedia document) arranged in a specific order. Thus the `cause` relation in figure 1, for example, could be realized with nucleus first, or satellite first, or satellite embedded within nucleus:

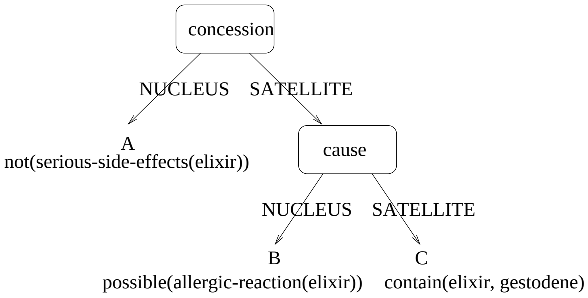

Figure 1: Rhetorical representation

(2a) Elixir occasionally provokes a mild allergic reaction^B, because it contains gestodene^C.

(2b) Because it contains gestodene^C, Elixir occasionally provokes a mild allergic reaction^B.

(2c) Elixir, because it contains gestodene^C, occasionally provokes a mild allergic reaction^B.

In the RAGS proposal, which aims to extract a useful common approach from current work in NLG, the `DocRep` comprises an ordered tree corresponding roughly to the 'logical markup' in notations like HTML and LaTeX. More precisely, a distinction is made between *abstract* and *concrete* levels of representation, where the abstract representation corresponds to logical markup (e.g., concepts like 'paragraph' and 'emphasis'), while the concrete representation also covers graphical markup (concepts like 'vertical space' and 'bold face'). In terms of this distinction, it is the `AbsDocRep` that is specified during text planning; graphical markup can be deferred to a later formatting stage.

Figure 2 shows two alternative document representations expressing the rhetorical content in figure 1. Following Power (2000), the nodes of the tree are labelled with 'text-categories' using a system that extends the 'text grammar' proposed by Nunberg (1990).[1] These document

[1] Nunberg's terms 'text-phrase', 'text-clause', and 'text-sentence' refer to *textual* categories, which should not be confused with their syntactic counterparts. They are defined not by syntactic formation rules but by their role in text-structure, which is typically marked as follows: *text-sentences* begin with a capital letter and end in a full stop; *text-clauses* are separated by semicolons; *text-phrases* are

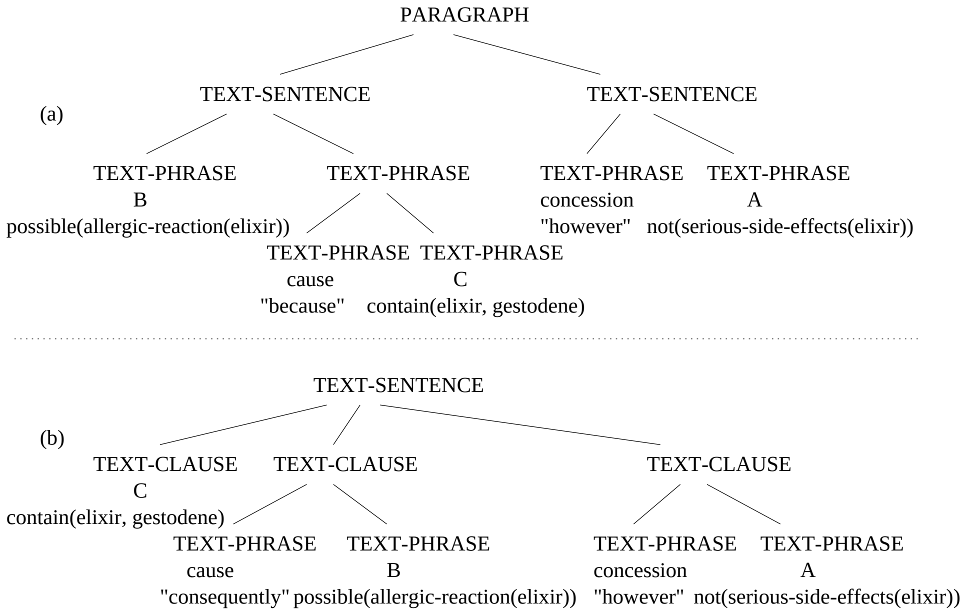

Figure 2: Document representations

representations can now be passed to the tactical generator for the syntactic realization of the elementary propositions; the resulting texts might be as follows:

(3a) Elixir occasionally provokes a mild allergic reaction^B, because it contains gestodene^C. However, Elixir has no serious side-effects^A.

(3b) Elixir contains gestodene^C; consequently, it occasionally provokes a mild allergic reaction^B; however, Elixir has no serious side-effects^A.

3 Structural compatibility

Summarising the argument so far, we have made three main points:

- Rhetorical structure has typically been represented by unordered RST trees such as figure 1.
- Document structure, which conveys information similar to logical markup in HTML, can be represented by ordered trees in which nodes are labelled with text-categories (figure 2).
- A given rhetorical representation can be expressed by a variety of different document representations, in which the propositions occur in different orders, and in different text-category configurations, and the rhetorical relations are expressed by different connectives.

This formulation of the problem raises an obvious question: how can we characterize the set of document representations that adequately realize a given rhetorical representation? Elsewhere (Power, 2000), we have argued that an adequate realization must meet three conditions:

Correct content:
All propositions and rhetorical relations must be expressed.

Well-formed structure:
General formation rules for document structure must be respected (e.g. a text-sentence cannot contain a paragraph, unless the paragraph is indented).

Structural compatibility:
The document representation must organize the propositions in a way that is compatible with their organization in rhetorical structure.

constituents of text-clauses, sometimes separated by commas, although within text-clauses the hierarchical structure is expressed mainly through syntax.

The first two conditions are relatively straightforward, but what is meant exactly by 'structural compatibility'?

Assuming that we are comparing two trees, the strongest notion of compatibility is *isomorphism*, which can be defined for our purposes as follows:

> `DocRep` is *isomorphic* with `RhetRep` if they group the elementary propositions in exactly the same way.
>
> *More formally, every set of propositions that is dominated by a node in* `DocRep` *should be dominated by a node in* `RhetRep`*, and vice-versa.*

Under this definition, the rhetorical representation in figure 1 is isomorphic with the document representation in figure 2a, but not with that in figure 2b:

- Proceeding top-down and left-to-right, the five nodes in figure 1 dominate the proposition sets $\{A, B, C\}$, $\{A\}$, $\{B, C\}$, $\{B\}$, and $\{C\}$.
- Ignoring nodes that express discourse connectives, the nodes in figure 2a dominate the proposition sets $\{A, B, C\}$, $\{B, C\}$, $\{B\}$, $\{C\}$ (twice), and $\{A\}$ (twice). These are exactly the same sets that were obtained for figure 1.
- The corresponding sets for figure 2b are $\{A, B, C\}$, $\{C\}$, $\{B\}$ (twice), and $\{A\}$ (twice). Since the set $\{B, C\}$ is missing from this list, there is a grouping in figure 1 that is not realized in figure 2b, so these representations are not isomorphic.

Since structures like figure 2b are common, isomorphism seems too strong a constraint; we have therefore proposed (Power, 2000) the following weaker notion of compatibility:

> `DocRep` is *compatible* with `RhetRep` if every grouping of the elementary propositions in `DocRep` is also found in `RhetRep`.
>
> *Formally, every set of propositions that is dominated by a node in* `DocRep` *should be dominated by a node in* `RhetRep` — *but the converse is not required.*

Under this constraint, we allow the document representation to omit rhetorical groupings, but not to introduce new ones. The resulting structures may be ambiguous, but this will not matter if the unexpressed rhetorical relationships can be inferred from the content.

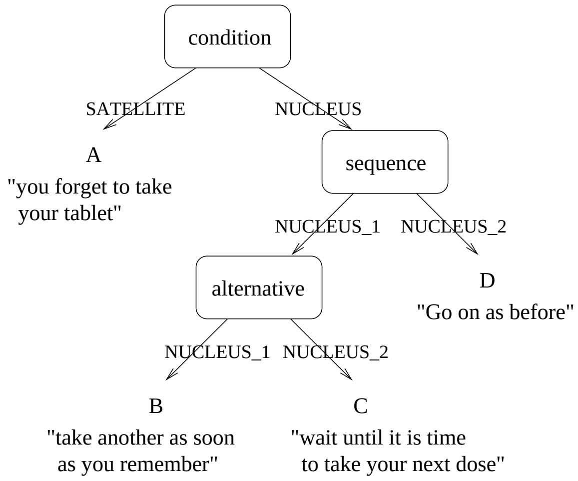

Figure 3: Rhetorical representation of instruction

4 Extraposition

The compatibility rule may be a useful text-planning heuristic, but as a constraint on adequacy it still seems too strong. Looking through our corpus of patient information leaflets, we have noticed some exceptions, especially in passages giving conditional instructions:

(4) If you forget to take your tablet^A, take another as soon as you remember^B or wait until it is time to take your next dose^C. Then go on as before^D.

From the point of view of document structure, this passage is a paragraph comprising two text-sentences: thus the proposition D is separated from the other three propositions, which are grouped in the first sentence. However, rhetorically speaking, D belongs to the consequent of the conditional: it is the final step of the plan that should be activated if the patient forgets to take a dose (figure 3). Compatibility is violated because the `DocRep` contains a node (the first text-sentence) dominating the proposition set $\{A, B, C\}$, which is not dominated by any node in figure 3.

Such examples might be explained as the result of loose punctuation or layout, perhaps through imitation of the patterns of conversation, in which extra material is often tagged on as an afterthought. Thus proposition D remains grouped with B and C — they occur consecutively — but through a minor relaxation of normal punctuation it has been separated by a full-stop rather than a comma. However, this explanation fails to cover variations of the example in which the propositions in the consequent are not realized consecutively in the `DocRep`:

(5) Consult your doctor immediately^A if a rash develops^B. It might become seriously infected^C.

In this example, A must be grouped rhetorically with C rather than with B, unless we take the radical step of allowing rhetorical structure to contradict logical structure. The proposition C cannot be logically conjoined with the conditional because it contains a hypothetical discourse referent (the rash) that is bound to the antecedent, and is therefore inaccessible outside the conditional.

If passages of this kind are not artifacts of loose punctuation, why do they occur? A plausible reason, we suggest, is that some complex rhetorical patterns cannot easily be realized in a way that maintains structural compatibility, usually because text-clauses are overloaded. Conditionals are especially prone to this problem because the only common discourse connective ('if') is a subordinating conjunction which can only link spans within a syntactic sentence (and thus within a text-clause). If either the antecedent or the consequent is complex, the author is faced with a tricky problem. We have found examples in patient information leaflets of conditional sentences so long that they are almost incomprehensible. More skilled authors, however, succeed in presenting the material clearly either by using layout (e.g., a complex antecedent is presented as an indented list), or by a trick of rhetorical reorganization that we will call *extraposition*. It is this trick that introduces an incompatibility between `RhetRep` and `DocRep`.

Extraposition typically occurs when a rhetorical representation R contains a complex embedded constituent C. To respect structural compatibility, R should be realized by a document unit that contains the realization of C; instead, in extraposition, a document unit realising $R - C$ is *coordinated* with one realizing C, so that the extraposed material C is raised in the `DocRep` to the same level as R. To reconstruct the meaning of the whole passage, the reader has to plug C back into R. In most cases, the author facilitates this task through an explicit deictic reference to the extraposed material (Bouayad-Agha et al., 2000):

(6) If you have any of the following, tell your doctor:

difficulty in breathing
abdominal pains
nausea or vomiting

Occasionally, however, the author leaves the extraposition implicit, assuming that the reader can infer the correct location of C within R from the propositional content. In such cases, the extraposition looks like an afterthought, because the unit realizing $R - C$ contains no signal that a gap in its content will be filled in later.

We have also come across rare examples of another kind of incompatibility in which Marcu's (1996) *principle of nuclearity* is violated by grouping together two satellites which have the same nucleus. Suppose that the rhetorical representation in figure 1 is realized by the following passage, in a context in which the reader knows nothing about gestodene:[2]

(7) Although Elixir has no serious side-effects^A, it contains gestodene^C. Consequently, it occasionally provokes a mild allergic reaction^B.

The apparent concession relation between A and C here is paradoxical, since in rhetorical structure they are unrelated. Of course a contrast between A and C might be perceived by a medical expert; however, one can construct similar examples in which the apparent relation is even less plausible:

(8a) Although we usually work from nine to five^A, today is Friday^C. Consequently, we can go home early^B.

This may be rather loose, but many people find it acceptable. It could be explained as a rhetorical trick in which the sheer paradox of the con-

[2]**Post-proceedings erratum:** Figure 1 is not the rhetorical representation of text (7). Instead, the rhetorical representation of (7) is one in which A is the satellite of B in the concession relation and C is the satellite of B in the evidence relation. We apologise for this error, although it does not alter the point of our argument.

cession serves as a signal that it is incomplete. The device might be spelled out as follows:

> Although Elixir has no serious side-effects^A, *there exists a contrasting state of affairs resulting from the fact that* it contains gestodene^C. *This state of affairs is that* it occasionally provokes a mild allergic reaction^B.

Unlike the conditional examples above, this device works only when the rhetorically grouped propositions B and C are consecutive in the `DocRep`. Thus whatever view is taken of example (8a) , everyone finds its variant (8b) much worse:

(8b) # Today is Friday^C although we usually work from nine to five^A. Consequently, we can go home early^B.

5 Implications for NLG

For many NLG applications, the notion of compatibility defined above is a useful hard constraint; even if violations of this constraint are sometimes acceptable, they are not essential. However, for some kinds of material (e.g., complex instructions), extraposition is a convenient rhetorical device which might improve the readability of the generated texts, so it is worth considering how a text planner might be configured so as to allow solutions that violate compatibility.

In terms of the RAGS framework, there are broadly two possible approaches. First, we could introduce incompatibility by defining *transformations* on the `RhetRep`; alternatively, we could relax the constraints governing the transition from `RhetRep` to `DocRep`. The RAGS proposal (1999) allows for rhetorical transformations through a distinction between abstract and concrete rhetorical representations. The abstract representation `AbsRhetRep` expresses the rhetorical content of the underlying message, while the concrete `RhetRep` expresses the rhetorical structure directly realized in the text and corresponds to the representation used by Scott and Souza (1990) to discuss textual realisation. If `RhetRep` is incompatible with `AbsRhetRep`, the text structure `DocRep` will also be incompatible with `AbsRhetRep`, even though the rules for realizing rhetorical structure by document structure are themselves compatibility-preserving. Transformation operations are also used by Marcu (2000) to map Japanese rhetorical structures onto English-like rhetorical structures, but these are mappings between two `RhetRep`s rather than from an `AbsRhetRep` to a `RhetRep`.

If transformations are allowed, there are obvious dangers that the message will be expressed in such a distorted way that the reader cannot recover the original intention. For this reason, rhetorical transformations must be defined with care. A fairly safe option would appear to be the extraposition of a proposition elaborating the antecedent of a conditional — even though such a transformation would violate Marcu's (1996) 'nuclearity principle' (assuming that the antecedent is regarded as the satellite). The following examples show that this transformation leads to acceptable texts regardless of the order of nucleus and satellite within the conditional:

(9a) Do not use Elixir if you have had an allergic reaction to Elixir. An allergic reaction may be recognised as a rash, itching or shortness of breath.

(9b) If you have had an allergic reaction to Elixir, do not use Elixir. An allergic reaction may be recognised as a rash, itching or shortness of breath.

However, the approach based on rhetorical transformations leads to difficulties when the acceptability of the resulting text depends on linear order as well as grouping. For instance, suppose that we try extraposing the elaboration of a satellite when the main relation is not a conditional, but a concession. The following passages show two texts that might result, but in this case the second version sounds anomalous: even if they are not grouped together in the `DocRep`, the satellite and its elaboration at least need to be consecutive.

(10a) You should not stop taking Elixir, even though you might experience some mild effects. For example, feelings of dizziness and nausea are very common at the beginning of treatment.

(10b) # Even though you might experience some mild effects at the beginning of the treatment, you should not stop taking Elixir. For example, feelings of dizziness and nausea are very common at the beginning of treatment.

A transformation from `AbsRhetRep` to `RhetRep` cannot distinguish these cases, so that

10a is allowed while 10b is prohibited, unless the `RhetRep` is at least partially specified for linear order. Adhering strictly to the RAGS framework, where linear order is specified only in `AbsDocRep`, one would have to adopt the alternative of building an incompatible `AbsDocRep` from `RhetRep`, constraining the linear order at this stage.

6 Conclusion

We have discussed various examples of extraposition. This phenomenon is due to various factors: the complexity of the material (example 4), the presence of logical information (5), the use of referring expressions to access information at various degrees of accessibility in the text structure (5,6,9), and the use of particular rhetorical strategies (7,8). This last group of examples concerns a concession construction similar to the one discussed by Grote et al. (1997), namely the *substitution concession*. This type of concession groups together the conceded part A and the explanation C but leaves the conclusion B unverbalised. The difference in the case of examples 7 and 8 is that A and C are grouped together but B is required to follow them because there is not enough information for the reader to infer B from A and C.

The extraposition phenomenon shows that the nucleus-satellite distinction is not the only factor influencing the segmentation of the message. In example 10, the injunction *you should not stop taking Elixir* obviously expresses the main intention of the author. However, the fact that the subordinated concession is placed after its main clause makes it available for further expansion. The sometimes competing informational and intentional roles of discourse segments have been at the centre of the debate over the nucleus-satellite distinction (Moore and Pollack, 1992; Moser and Moore, 1996; Bateman and Rondhius, 1997); the accessibility of discourse segments on the right frontier of a discourse structure is a phenomenon that has already been discussed by several researchers (Webber, 1991; Asher, 1993). Extraposition provides a useful and sometimes important means of rearranging complex material in an abstract discourse representation in order to satisfy the constraints posed by linearisation into text.

References

N. Asher. 1993. *Reference to Abstract Objects in Discourse.* Kluwer Academic Publishers, Netherlands.

J. Bateman and K. Rondhius. 1997. Coherence relations: Towards a general specification. *Discourse Processes*, 24(1):3–50.

N. Bouayad-Agha, D. Scott, and R. Power. 2000. Integrating content and style in documents: a case study of patient information leaflets. *Information Design Journal*, 9(2):161–176.

B. Grote, N. Lenke, and Stede M. 1997. Ma(r)king concessions in english and german. *Discourse Processes*, 24(1):87–117.

W. Mann and S. Thompson. 1988. Rhetorical structure theory: towards a functional theory of text organization. *Text*, 8(3):243–281.

D. Marcu, L. Carlson, and M. Watanabe. 2000. The automatic translation of discourse structures. In *Proceedings of the North American Chapter of the Association for Computational Linguistics (NAACL'2000)*, Seattle, Washington.

D. Marcu. 1996. Building up rhetorical structure trees. In *Proceedings of AAAI-96*. American Association for Artificial Intelligence.

D.J Moore and M.E. Pollack. 1992. A problem for rst: The need for multi-level discourse analysis. *Computational Linguistics*, 18(4):537–544.

M. Moser and J.D. Moore. 1996. Towards a synthesis of two accounts of discourse structure. *Computational Linguistics*, 22(3):409–419.

G. Nunberg. 1990. *The Linguistics of Punctuation.* CSLI, Stanford, USA.

R. Power. 2000. Mapping rhetorical structures to text structures by constraint satisfaction. Technical report, ITRI, University of Brighton.

RAGS. 1999. The RAGS project: towards a reference architecture for natural language generation systems. Technical report, Information Technology Research Institute, Brighton, UK.

D. Scott and C. de Souza. 1990. Getting the message across in RST-based text generation. In R. Dale, C. Mellish, and M. Zock, editors, *Current Research in Natural Language Generation.* Cognitive Science Series, Academic Press.

B.L Webber. 1991. Structure and ostension in the interpretation of discourse deixis. *Language and Cognitive Processes*, 6(2):107–135.