Beyond elaboration: generating descriptive texts containing it-clefts

Jon Oberlander Alistair Knott Michael O'Donnell Chris Mellish

January 29, 1999

Abstract

Algorithms for discourse processing commonly assume that texts are tree-structured, and that the coherence of a constituent within a discourse tree can be explained locally, in terms of its subconstituent spans. We argue that this notion needs to be qualified in various ways. Taking RST as a starting point, we examine several types of non-local dependency, all involving the rhetorical relation known as (object-attribute) ELABORATION. We suggest that this relation is of a different type from the others in the RST canon, and that the phenomena it is used to cover are better explained in terms of local and global focus. We propose a model distinguishing **entity-based** relations from conventional rhetorical relations, and allowing non-local dependencies to be expressed by relations of this type. We show how this model is well-suited for a treatment of *it*-clefts in English.

1 Introduction

It is widely assumed that coherent extended texts have hierarchical structure. In the key computationally-oriented theories of discourse structure—Grosz and Sidner (1986); Mann and Thompson (1988); Hobbs (1985)—this assumption takes its simplest form: coherent texts are held to be uniformly and recursively hierarchical, thus possessing a tree-shaped structure. A text planning system based on this model of discourse can be formulated as a tree-building algorithm.

Meanwhile, English possesses a family of cleft constructions. The cleft involves a departure from canonical SVO order because it extracts a constituent from its normal position. Pragmatic effects can arise from the use of such non-canonical structures, and it follows that the ability to produce such structures would, in principle, be a useful behaviour in a text planning system.

In this paper, we examine some problems with a hierarchical approach to text structure,

DRAFT

and then discuss issues in the automatic generation of clefts by way of illustration. We propose a theory of discourse organisation which escapes the problems we have outlined, and indicate how it would lead to the generation of cleft constructions in appropriate circumstances.

2 RST and some assumptions about text structure

It is widely assumed that coherent texts are uniformly and recursively hierarchical, possessing a tree-shaped structure. On this model, text planning systems can be formulated as tree-building algorithms. The trees built are sometimes of rhetorical relations, sometimes of more complex schemas, but in any case they permit a recursive approach to planning, whereby decisions about the composition of a constituent in the text structure are **local** to that constituent. We focus here on planning approaches founded on Mann and Thompson's (1988) Rhetorical Structure Theory (RST), according to which (with some simplifications) a text can be modelled as a binary branching tree of **text spans**, where each pair of sister spans is linked by a rhetorical relation. This theory permits three central assumptions:

- Nuclearity: a complex span (comprising a nucleus and a satellite) can be linked to another span via a rhetorical relation iff its nucleus span can be so linked.
- Continuous constituency: the nucleus and satellite spans of a relation appear adjacent to one another in the text.
- Non-re-entrancy: each span is related to exactly one other span.

These assumptions do indeed license well-structured texts. We illustrate with part of a text produced by Ilex-2, a generation system which delivers a sequence of descriptions of artefacts in a tour of a museum gallery:

(1) (1) This jewel draws on natural themes for inspiration; (2) it is a remarkably fluid piece. (3) Indeed, Organic style jewels usually draw on natural themes for inspiration; (4) for instance the organic brooch we saw earlier looked crystalline.

The structure for this text is given in Figure 1. By the nuclearity assumption, the top-level AMPLIFICATION relation holds between the complex spans (1–2) and (3–4) in virtue of their

DRAFT DRAFT DRAFT DRAFT DRAFT DRAFT DRAFT DRAFT DRAFT DRAFT DRAFT DRAFT DRAFT DRAFT DRAFT DRAFT DRAFT

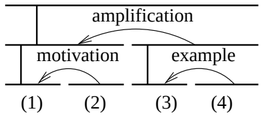

Figure 1: RST Analysis of Example 1

respective nuclear spans, (1) and (3). The expansions of (1) with (2), and of (3) with (4), take place independently of the higher-level relation. By continuous constituency, satellite spans appear adjacent to their nuclei; and by non-re-entrancy, each span is linked to just one other span. Here, adherence to these assumptions results in a well-structured text.

3 Some structural problems with ELABORATION

While these assumptions are useful for a text planner, they are also problematic in a number of respects, as has often been noted: see for instance Mooney *et al* (1990), Sibun (1992) and Kittredge *et al* (1991). Our central concern in this paper is to associate these problems with one relation in particular, namely the RST relation known as (object-attribute) ELABORATION.[1] Mann and Thompson define this to hold between two spans if the nucleus 'presents' an object (i.e. contains a mention of it) and the satellite subsequently presents an attribute of that object. The precise meaning of 'attribute' is not clear, but the relation is intended to have a very wide interpretation: allowing two spans to be linked whenever they are both 'about' the same entity. In the type of text which our system produces—a series of descriptions of a collection of related entities—this relation is heavily applicable, and the problems we note are thus quite widespread.

3.1 Discontinuous constituency

An initial problem is illustrated in the following text, taken from a museum guidebook.

[1]Other types of ELABORATION—for instance, what Mann and Thompson call 'process step' elaboration or 'generalisation-specific' elaboration—do really apply between propositions. In what follows, references to ELABORATION are exclusively to the object-attribute variety.

DRAFT DRAFT DRAFT DRAFT DRAFT DRAFT DRAFT DRAFT DRAFT DRAFT DRAFT DRAFT DRAFT DRAFT DRAFT DRAFT DRAFT

(2) (1) In the women's quarters the business of running the household took place. (2) Much of the furniture was made up of chests arranged vertically in matching pairs (...). (3) Female guests were entertained in these rooms, which often had beautifully crafted wooden toilet boxes with fold-away mirrors and sewing boxes, and folding screens, painted with birds and flowers.

(4) Chests were used for the storage of clothes...

In this text, an entity mentioned in the middle of the first paragraph, *chests*, becomes the focus of the second paragraph. We can refer to this move pre-theoretically as a **resumption**.[2]

The move is clearly legitimate in the above context, and yet an analysis in terms of a tree of relations is difficult. The problem is that sentence 4 needs to be seen as the satellite of an ELABORATION relation, but the obvious nucleus for this relation—sentence 2—is not accessible; if we analyse sentences 2 and 3 as ELABORATIONs of sentence 1, as seems necessary, we have effectively closed off sentence 2 as the nucleus for further ELABORATIONs. In order to treat sentence 4 as an ELABORATION of sentence 2, we would have to analyse sentence 3 as being subordinate to sentence 2: this analysis seems inappropriate; moreover, it makes the position of the paragraph break hard to explain. One other option is to consider sentences 2 and 3 to form a single span, but such an approach would sacrifice apparent structural detail. Note that we cannot just ignore the relationship between sentences 2 and 4 in our representation of the text: it is only because the chests are mentioned in the former sentence that they are a relevant topic for discussion. To account for coherence in this case, it seems we must either abandon compositionality, in some circumstances, or adopt a notion of discontinuous constituency for text spans.

A particularly common manifestation of this problem is in cases of parallelism within discourse structure. Especially in descriptive texts, it is common for a number of entities to be

[2]The notion of a resumption bears some resemblance to Grosz and Sidner's notion of a **digression**. This is a discourse segment which (a) is not related to the immediately preceding segment by dominance or satisfaction-precedence, and (b) contains mention of an entity salient in the interrupted segment. However, Grosz and Sidner's definition covers cases where there is deviation from the (possibly joint) discourse plan; furthermore, the definition implies that a link due to a common entity can only occur between *adjacent* segments; our claim is that resumptions can occur between non-adjacent segments.

DRAFT DRAFT DRAFT DRAFT DRAFT DRAFT DRAFT DRAFT DRAFT DRAFT DRAFT DRAFT DRAFT DRAFT DRAFT DRAFT DRAFT DRAFT

introduced sequentially in a sequence of spans, and then elaborated on in subsequent spans in the order of their introduction. Accounting for these subsequent mentions as ELABORATIONs of the spans where they were introduced is not possible without violating adjacency or compositionality constraints. Mann and Thompson acknowledge from the outset that RST cannot account for the constraints which apply in such contexts. McKeown (1985) deals extensively with cases of parallelism in text, although this account is not set in the context of a theory of coherence relations. Kittredge *et al* (1991) give several examples of parallelism; indeed, in one case they identify ELABORATION as the relation responsible for the problem.

3.2 Nuclearity and embedding

The preceding section presents a case where a context-free account of relations undergenerates the space of possible texts. There are also cases where it overgenerates; again, these relate principally to the ELABORATION relation. There often seem to be difficulties in embedding ELABORATIONs within other relations. Consider:

(3) (1) Arts-and-Crafts jewels tend to be elaborate. (2) However, this jewel has a simple form.

This text contains a CONCESSION relation whose nucleus is 2 and whose satellite is 1. In principle, we could expand either span with additional relations. But note what happens when we embed an ELABORATION under span 1:

(4) (1) Arts-and-Crafts jewels tend to be elaborate. (1a) They are often made by skilled single craftsmen. (2) However, this jewel is simple in form.

Sentence 1a elaborates on 1 by providing more information about Arts-and-Crafts jewels. However, it also makes it hard to attach sentence 2 to sentence 1. Note that there *is* a coherent interpretation of the text, if 1a is treated as somehow expanding on the fact that Arts-and-Crafts jewels are elaborate, rather than simply as presenting more information about the jewels. For example, it might be that skilled individuals are more likely to be able to turn out intricate designs, than are teams of semi-skilled artisans, or the machines then available for

DRAFT DRAFT DRAFT DRAFT DRAFT DRAFT DRAFT DRAFT DRAFT DRAFT DRAFT DRAFT DRAFT DRAFT DRAFT DRAFT DRAFT

mass production. But then 1a offers a reason for 1, and on this interpretation, the embedded relation is no longer object-attribute ELABORATION—and that is precisely our point.

Note also that the problem is not just due to difficulties with 'high-level' relations in general, or with 'left-branching' tree structures. Compare an embedded EXAMPLE relation:

(5) (1) Arts-and-Crafts jewels tend to be elaborate. (1a) For instance, the previous jewel had ornate festoons. (2) However, this jewel is simple in form.

It is true that there may be a limit to the depth of embedding permissible for any relation, particularly for left-branching RS trees. It is also true that we have not shown conclusively that ELABORATION can never be embedded. However, the point here is that ELABORATION does appear to resist even the simplest kind of embedding, and it does so where other relations can happily be embedded. This provides a further reason to suspect that it is qualitatively different from the other relations.

3.3 Attributehood and information structure

A third problem is that the definition of ELABORATION is extremely vague. Mann and Thompson do not formally define the notions of 'object' and 'attribute' in object-attribute ELABORATION; however, there seem good reasons for distinguishing between a sentence which presents an 'attribute' of an entity, and one which simply mentions it. Consider:

(6) Jessie King was a silversmith. She worked mainly in London.

(7) Jessie King was a silversmith. ?She designed a silver ring.

While the sentences in each text have an entity in common (Jessie King), the second sentence of Example 7 seems at best an inconsequential continuation.[3] We can imagine any number of facts 'involving' Jessie King, but only a subset of these (her place of work, date of birth, etc) are relevant here as part of a description of her.

Note that we cannot account for attributehood in terms of linguistic structures like subject and predicate; for instance, *King* appears in both of the above texts in subject posi-

[3] Unless we imagine that designing a silver ring is somehow a noteworthy achievement for a jewellery designer. We are assuming this is not the case.

DRAFT DRAFT DRAFT DRAFT DRAFT DRAFT DRAFT DRAFT DRAFT DRAFT DRAFT DRAFT DRAFT DRAFT DRAFT DRAFT DRAFT DRAFT

tion. Equally, the relation between attributehood and the information-structural notions of topic and comment is not immediately clear. Take topic and comment as used by Gundel (1985,1988). A topic is an entity that is in some way familiar to both speaker and hearer. Thus, an entity will be the topic of a sentence if the speaker intends to alter the hearer's knowledge or intentions concerning that entity. A predication will be a sentence's comment if the speaker intends the predication to be assessed relative to the topic. Perhaps Example 7 is odd because *designed a silver ring* does not provide an appropriate comment for the topic supplied by *King*. Similarly, on van Kuppevelt's question view of topics (1995), if the topic here is supplied by the question *Who is Jessie King?*, the answer supplied in Example 7 is less satisfactory than that in Example 6. And on any view of topic, if the ring is actually supposed to be the topic, the syntactic structure does not indicate this.

At the very least, it seems that to make sense of attributehood, we must actually make reference to structures in the underlying knowledge base itself, where facts can be indexed according to the entities they are primarily about.

3.4 Re-entrancy and clefts

A final problem with ELABORATION follows from this. Consider the following variant on Example 7:

(8) Jessie King was a silversmith. She designed the silver ring we saw earlier.

This text is a significant improvement on Example 7. This is apparently because the object of the second sentence is an entity which must have already been mentioned in the discourse. We call such structures **joint focus returns**. They are often very useful in our domain, allowing the inclusion of facts omitted from earlier descriptions due to space limitations, and the reiteration of previously-mentioned facts which may have been forgotten.

Again, note that a joint focus return is a *non-local* phenomenon. To define it, we need to make reference to a point arbitrarily far back in the text. What is more, it results in an element of *re-entrancy* in the text structure: in order to account for a joint focus return, it must be related both to the sentence which introduces it, *and* to the original reference to

DRAFT DRAFT DRAFT DRAFT DRAFT DRAFT DRAFT DRAFT DRAFT DRAFT DRAFT DRAFT DRAFT DRAFT DRAFT DRAFT DRAFT

the discourse-old entity. Trying to account for a phenomenon such as this within RST would involve further violations of the assumptions it stipulates.

Note finally that re-entrant structures such as that in Example 8 are especially suitable for realisation using cleft structures:

(9) Jessie King was a silversmith. It was she who designed the silver ring we saw earlier.

The question of when such syntactic structures are possible in a discourse is one which has received much attention; see e.g. Gazdar (1979), Collins (1991), Delin and Oberlander (1995). However, in the light of examples such as 9, it seems doubtful that an account of structures such as clefts can be framed within a theory which enforces a rigid tree structure on a text.

4 Beyond ELABORATION in RST

All the problems noted above are traceable in some way to object-attribute ELABORATION. At best, they indicate that this relation constitutes an exception to the attractive initial assumptions of context-independence. However, an alternative possibility is simply that ELABORATION should be omitted entirely from the set of relations. While several commentators have noted problems with ELABORATION, none have considered what an account of discourse relations would look like without this relation: it is this question which we would like to address.

A crucial point to note in this regard is that the discourse phenomena described by ELABORATION overlap extensively with phenomena described by other theories of discourse, namely those concerned with focus structure. Consider firstly theories of local focus, such as Sidner's (1981). In particular, take Grosz *et al*'s (1995) account of centering. A primary concern for this theory is to catalogue the different discourse structures which can obtain in cases where two adjacent sentences make reference to a common entity. The issue is explored both in hypotheses about how this entity should be referred to in the second sentence (for example, pronominally) and about which sentence configurations make for 'good continuations'. The centering account is explicitly entity-based, and is expressed at a level of detail far greater than that given in the definition of ELABORATION, which *prima facie* covers the same cases.

DRAFT DRAFT DRAFT DRAFT DRAFT DRAFT DRAFT DRAFT DRAFT DRAFT DRAFT DRAFT DRAFT DRAFT DRAFT DRAFT DRAFT DRAFT

Moreover, it is not bound by the hierarchical constraints imposed on RST relations which were shown to be problematic for ELABORATION: adjacent sentences are related in chains, rather than in trees.

ELABORATIONs at a higher level of hierarchy are also problematic. As Example 2 indicates, the set of entities in a complex text span to which resumptions can be made includes some in apparently inaccessible positions. By RST's compositional assumptions, a complex text span is ultimately connected to surrounding text *via a single proposition*—typically, the initial one in the span. Cases of resumption cast doubt on this assumption. Indeed, it seems odd, given a complex span of text 'about' one entity, to accord a single proposition special status. It is better, both empirically and intuitively, to take an entity to be in (global) focus in the span, and to consider any other entities introduced in the span as potential foci for resumption.

ELABORATION has frequently been treated in the literature as a special kind of relation. For instance, Marcu's (1997) algorithm for identifying the relations in a text from surface cues relies exclusively on discourse markers for all relations except ELABORATION (and JOINT); for these latter two relations, word co-occurrence measures provide the strongest surface indicators. Similarly, Moser *et al* (1996) identify a class of (informational) relations termed 'subrelatum' relations, which apply between components of propositions (typically entities) rather than whole propositions: varieties of ELABORATION account for most of the relations in this category. These peculiarities of the ELABORATION relation, taken together with the structural problems attributed to it in the preceding sections, suggest an account in which ELABORATION is subject to distinctive structural constraints.

We can begin by noting some requirements for this account. Firstly, it must set out an adequate notion of attributehood: in particular, one which allows the distribution of cleft structures to be charted. Secondly, it must allow the kind of non-local dependencies noted in the previous section; in particular, constraints concerned with parallelism must be expressible. Thirdly, it must interact productively and non-redundantly with the (now partial) relation-based account of discourse structure.

DRAFT DRAFT DRAFT DRAFT DRAFT DRAFT DRAFT DRAFT DRAFT DRAFT DRAFT DRAFT DRAFT DRAFT DRAFT DRAFT DRAFT

5 Background: *It*-clefts in discourse

Before outlining our proposed account, we should provide some more detail on the kinds of cleft construction we aim to capture. We have focussed in particular on cleft constructions in English. There follows an overview of the relevant phenomena.

The cleft in English takes several forms; given the canonical form *John ate the beans*, we can re-express this content using different types of cleft:

WH-cleft What John ate was the beans.

Reverse WH-cleft The beans were what John ate.

***It*-cleft** It was John who ate the beans.

The initial constituent of a cleft has been termed in the literature the CLEFTED CONSTITUENT (sometimes known as the cleft's FOCUS). The relative-clause-like post-copular element has been referred to as the CLEFT CLAUSE.

In the above cases, use of the cleft construction appears to evoke some logical presupposition not present in the canonical case (cf. Gazdar, 1979, Collins, 1991). Of course, such effects are usually most apparent in extended discourses, and we can easily isolate discourse contexts in which a cleft appears appropriate, whereas the related canonical construction appears infelicitous. Compare the contrastive (10), from the LOB corpus, and its constructed canonical equivalent (11) (Delin and Oberlander, 1995's examples [7] and [8]):

(10) Doubling the selling space to 700 square feet was not to be the greatest expense. *It was the new fixtures and fittings to fill this space that would be costly.*

(11) ?Doubling the selling space to 700 square feet was not to be the greatest expense. *The new fixtures and fittings to fill this space would be costly.*

It-clefts are reasonably common in written text, and WH and reverse-WH clefts are particularly common in speech. Collins (1991) gives the following general frequencies per 10,000 words: in written texts, *it*-clefts occur 5.7 times, and WH-clefts occur 4.1 times; in speech, *it*-clefts occur 4.3 times, and WH-clefts occur 14.3 times (p. 181). We have been examining

DRAFT DRAFT DRAFT DRAFT DRAFT DRAFT DRAFT DRAFT DRAFT DRAFT DRAFT DRAFT DRAFT DRAFT DRAFT DRAFT DRAFT

the genre of museum discourses—both written and spoken—and although we have not yet quantified the frequencies of occurrence, examples are easy to find.

5.1 Two types of *it*-cleft in discourse

Prince (1978) distinguishes two sub-types of *it*-cleft, which she termed stressed focus, and informative presupposition. Hedberg (1990) used the information-structural terms TOPIC-CLAUSE and COMMENT-CLAUSE respectively, and we follow this terminology here. Her notions of topic and comment follow that of Gundel introduced earlier. It should be noted that distinguishing the sub-types is primarily a way of outlining two different functions which the single construction can fulfill. However, as we will see, there are characteristic surface differences between clefts of the two sub-types. This is not entirely surprising, particularly in view of historical evidence that one sub-type may be actually be descended primarily from the reverse *wh*-cleft with similar information-structure (cf. Ball 1994).

In the first sub-type of *it*-cleft, information is presented in comment-topic order: topic clause clefts carry the topic in the cleft clause, and the comment in the clefted constituent. The clefted constituent is therefore the location of the new, frequently contrastive, information; the cleft clause is associated with a presupposition that is 'discourse-old' (in Prince, 1992's terms). Here is an example from a museum text:

(12) High-fired ceramics were first made in China at kilns in the coastal provinces of Jiangsu and Zhejiang as early as the Shang dynasty (c1600-c1050bc). However, *it was not until the Sui (ad 589-618) and Tang (618-906) dynasties that the Xing kilns in Hebei province and the Gongxian kilns in Henan, both in northern China, produced a translucent white ware which can be said to be the first porcelain.* [Wilkinson and Pearce (1996), p.51]

In the second sub-type, information is presented in topic-comment order: comment clause clefts carry the comment in the cleft clause, and the topic in the clefted constituent. The clefted constituent is short and generally anaphoric, while the subsequent cleft clause usually encodes a substantial amount of information discourse-new information (and since the cleft clause carries the sentence's logical presupposition, the presupposition is actually new to the

DRAFT DRAFT DRAFT DRAFT DRAFT DRAFT DRAFT DRAFT DRAFT DRAFT DRAFT DRAFT DRAFT DRAFT DRAFT DRAFT DRAFT DRAFT

discourse, and must presumably be accommodated, rather than matched to prior structure). Here is an example from the same museum text:

(13) The varieties of cotton plant are many, and are found throughout the world, particularly in places with a warm humid climate. Cotton cloth was used by the Aztecs of Mexico in the fifteenth century and some of the best cotton grows in the Nile valley. In India cotton was known as early as the third millennium BC when it was grown in the Indus valley.

It was cotton textiles from India that played an important part in the introduction of cotton to Japan by Portuguese merchants in the sixteenth century. By the late fifteenth century traders from Portugal had established trading posts in India; they were joined by the Dutch and English. [Wilkinson and Pearce (1996), p.93]

5.2 Clefts and discourse relations

Delin and Oberlander (1995) (pp. 483 ff.) propose a model of clefts in discourse which relates information structure to discourse coherence relations, via a model of presupposition processing, and temporal reference resolution. For current purposes, the relevant point is just this: the two sub-types of *it*-cleft possess differing information structures, and this ramifies through to the discourse coherence relations which they are supposed to enter into.

Topic-clause clefts can stand in QUESTION-ANSWER or in CONTRAST relations to prior discourse, like (12), above.[4]. Delin and Oberlander (1995) state:

> Both the contrast and the question-answer relation have in common that the topic-clause cleft provides the *completion* of a discourse segment, effectively closing off the dominating topic node and making it inaccessible for the building of further structure. [p484]

By contrast, it is claimed that comment-clause clefts stand in various kinds of BACKGROUND relation to prior discourse, like (13), above. They state:

[4]In this connection, it is worth noting that van Kuppevelt has argued for an alternative model, most appropriate to topic-clause clefts, which exploits the question-oriented view of topic mentioned earlier (cf. van Kuppevelt (1995))

DRAFT DRAFT DRAFT DRAFT DRAFT DRAFT DRAFT DRAFT DRAFT DRAFT DRAFT DRAFT DRAFT DRAFT DRAFT DRAFT DRAFT DRAFT

> [Rather] than closing off the segment to which it attaches, the comment-clause cleft leaves the node to which it attaches open for further elaboration.

Now, our comments above on ELABORATION shed some new light on the interaction between relations and clefts. For one thing, we can note a type of discourse context for clefts not found in Delin and Oberlander. Consider again the structure of a joint focus return:

(14) Jessie King was a silversmith. It was she who designed the silver ring we saw earlier.

Is this a topic-clause or a comment-clause cleft? It is slightly unusual, in that *both* the clefted constituent and the cleft clause present discourse-old information. However, Oberlander and Delin 1996 suggest that if the cleft-clause presents information merely inferrable from the discourse context, then it can be treated as topical. Because all the information presented is old, a case such as the one here can be considered 'all-topic'; nonetheless, it functions as a topic-clause cleft. This being so, note that the relation in question in this context is not one of those discussed by Delin and Oberlander 1995. That is, it seems to be neither CONTRAST nor QUESTION-ANSWER. Or at least, there is no *explicit* question being answered; if there is an implicit question, it could be *Who is Jessie King?* This is, effectively, a request for the value of an attribute of an object just introduced in the prior discourse. Thus, the only relation we could use in this case turns out to be ELABORATION, with its attendant problems. So we should at least add this relation to the contexts in which topic-clause clefts can appear.

Once ELABORATION has been identified as a context in which a topic-clause cleft can appear, we are led to ask whether we require adjacency between the span which introduces the cleft and the cleft itself. In simple cases of ELABORATION, as we have seen, the adjacency requirement can apparently be overridden. Does this also happen with clefts? The answer seems to be yes. Consider the following example from the LOB corpus:

(15) Mr. and Mrs. Kennedy received a particularly hearty welcome from students outside the Sorbonne University – where Mrs. Kennedy once studied. (. . .) At the airport, Mr. Kennedy praised his host as "a captain in the field in the defence of the West" for over 20 years (. . .).

It was Mrs. Kennedy who drew the crowds, said police.

DRAFT DRAFT DRAFT DRAFT DRAFT DRAFT DRAFT DRAFT DRAFT DRAFT DRAFT DRAFT DRAFT DRAFT DRAFT DRAFT DRAFT

Here, the clefted sentence refers to two discourse-old entities: Mrs Kennedy, and the crowds (who are inferrable from the hearty welcome). The point is that for both of these entities there is intervening material—reducing accessibility—between their first mention and their appearance in the cleft; in either case, the clefted sentence would be hard to analyse as an ELABORATION on a previous span without relaxing assumptions about adjacency or compositionality.

A final point relates to the function of a cleft as opening or closing a discourse segment. Delin and Oberlander suggested that topic-clause clefts tended to close segments, while comment-clause clefts tended to open them up. We can now propose a refinement of this idea, and suggest that a topic-clause cleft is bad as a locus for an entity-based resumption, while comment-clause clefts are legitimate in such a role. Note, however, that both kinds of cleft can serve as the nucleus in a tree of relations *if* ELABORATION *is removed from the set.*

5.3 Summary

We can draw a number of conclusions from the above discussion. To recap: our idea is that certain focus-moves are sufficient in themselves to ensure coherence in a text (joint focus returns being one such move), and that this allows us to drop ELABORATION from the set of relations. However, we should note that even if ELABORATION is removed from the set of relations, there are still cases where clefts appear within coherence relations. In Example 12, for instance, the nucleus of a CONCESSION relation is realised as a cleft. Indeed, examples of clefts within relations are not hard to find. What this suggests is that focus-based and relation-based moves are executed relatively independently in a discourse, with the only constraint being that every span should be introduced by at least *one* such move.

6 An account of entity-based and proposition-based coherence

We now proceed to a statement of the ELABORATION-less theory of descriptive text structure. We would like to preserve as much as possible of the RST-based model, and the hierarchical structuring algorithms it sanctions, while taking account of the exceptions due to ELABORATION noted above. It is true that some existing frameworks may be adequate for at least some

DRAFT DRAFT DRAFT DRAFT DRAFT DRAFT DRAFT DRAFT DRAFT DRAFT DRAFT DRAFT DRAFT DRAFT DRAFT DRAFT DRAFT

of our purposes: Halliday and Hasan (1976), for instance, distinguish referential relatedness from semantic connection within their three-way categorisation of different types of cohesion in text (p. 304). However, we aim for a slightly simpler account, and one closer to RST. Accordingly, we propose an account of text structure in which two types of coherence are distinguished: **proposition-based** coherence, which is defined according to the conventional constraints of RST (minus the ELABORATION relation), and **entity-based** coherence, which exists between spans of text in virtue of shared entities, and operates according to a different set of constraints.

The reader not interested in the details of this account can skip directly to example 16, which immediately precedes the conclusion to the paper.

6.1 Arg1s and Arg2s: indexing facts in the knowledge base

As noted earlier, one requirement for the account is a reference to structures in a knowledge base which differentiate a given proposition to be identified as being 'about' one or more of the entities it involves. (We can think of these as operating at the level of file cards, in Heim's (1982) terms, as extended by Vallduví, 1992.) The structures can be motivated in terms of considerations about the efficient indexing of information. For instance, we could represent the proposition *King designed R1* as a fact about *R1* or as a fact about *King*, or as both. The choice should depend on several factors: for instance, the fact that the mapping from designers to objects is one-to-many. For present purposes, we will assume a KB in which a fact involving several entities is indexed by only one of them, which we term its **Arg1**, and where secondary entities are termed **Arg2**, **Arg3** etc. We can then specify that a sentence 'presents an attribute' of an entity E iff it conveys a proposition whose Arg1 is E. If we assume that the Arg1 of the proposition *King designed the ring* is *the ring*, we thereby explain the difference between (6) and (7).

6.2 The content potential

The model of discourse we propose is formalised with reference to a structure called the **content potential**. This structure can be viewed as a resource for the tasks of content

DRAFT DRAFT DRAFT DRAFT DRAFT DRAFT DRAFT DRAFT DRAFT DRAFT DRAFT DRAFT DRAFT DRAFT DRAFT DRAFT DRAFT

selection and content organisation in natural language generation. The content potential is a graph, with three types of node (see Figure 2).

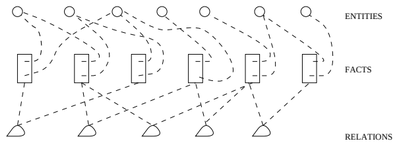

Figure 2: The form of the content potential

Entity-nodes represent entities, either individual (such as *this jewel*) or generic (such as *Art-Deco jewels*).

Fact-nodes represent facts about entities. We have adopted a simple model, in which each fact-node is represented by a **Pred** or predicate, applying between an **Arg1** and an **Arg2**, which are both entities, and are defined as in Section 6.1 above.[5] A fact-node also has a **Predarg** or 'predicate argument', which is a generic entity formed by abstracting away from the Arg1. For instance, the Predarg of the fact *This jewel was made by King* is the generic entity *jewels made by King*.

Relation-nodes represent possible coherence relations between pairs of fact-nodes. Each relation-node has a **Nuc** role and a **Sat** role, each of which is filled by a fact-node. A relation-node can also be linked to a set of other fact nodes called its **preconds**, which must be assimilated before the relation is presented. For instance, a CONCESSION relation of the form *P1, but P2* is linked to a fact node expressing that *if P1, then normally not P2*.

The space of possible texts creatable from a given initial entity or fact node can now be expressed in terms of constraints on graph traversal on the content potential. Where $F_{e1,e2}$ denotes a fact-node whose Arg1 is $e1$ and whose Arg2 or Predarg is $e2$, we define the following possible moves between one fact-node $F_{e1,e2}$ to another fact-node $F_{e3,e4}$, given a prior discourse D:

- A **relation-based** move $rel(F_{e1,e2}, F_{e3,e4})$ is possible iff there is a relation-node whose Nuc is $F_{e1,e2}$ and whose Sat is $F_{e3,e4}$.

[5] We are currently only considering binary Preds. Note that for sentences without an object, the Arg2 is taken to be a quality, such as *redness*.

DRAFT DRAFT DRAFT DRAFT DRAFT DRAFT DRAFT DRAFT DRAFT DRAFT DRAFT DRAFT DRAFT DRAFT DRAFT DRAFT DRAFT DRAFT

- An **Arg1-link** $A1(F_{e1,e2}, F_{e3,e4})$ is possible iff $e1 = e3$.
- An **Arg2-link** $A2(F_{e1,e2}, F_{e3,e4})$ is possible iff $e2 = e3$.
- A **joint focus return** $jfr(F_{e1,e2}, F_{e3,e4})$ is possible iff $e4 \in \{e1, e2\}$, and $e3$ is discourse-old. (An entity e is discourse-old given a prior discourse D if there exists a fact-node $F_{x,y}$ in D such that $e \in \{x, y\}$.)

(Note that the definitions of Arg1/Arg2-links are similar, though not identical to, those of continuation and shift in centering.)

6.3 Complex moves between fact-nodes

We now define the notion of a discourse **unit**.

- A **unit** is a fact-node or an **rs-tree**.
- An rs-tree is a tree of fact-nodes linked by relation-based moves, in which no fact-node occurs more than once.
- The fact-node at the root of an rs-tree is its **topnuc**. By convention, the topnuc of a fact-node is itself.
- Any Arg1-links, Arg2-links or joint focus returns between the topnucs of two units are taken to hold between the units too.
- An **entity-chain** is a sequence of units $U_1 \ldots U_n$ linked by Arg1-links, optionally fronted by a unit U_0, linked to U_1 by an Arg2-link, and optionally ended by a unit U_{n+1} linked to U_n by a joint focus return. No fact-node can occur more than once within the chain.
- The **focus** of an entity-chain is the Arg1 of its initial fact-node; its **expansion set** is the set of Arg1s, Arg2s and Predargs of all its fact-nodes.
- A **resumption relation** $Res(C_1, C_2)$ holds between two entity-chains C_1 and C_2 iff the focus F of C_2 is in the expansion set of C_1. In this case, we define the fact-node which **introduces** C_2 as the most recent fact-node in C_1 which features F as its Arg1, Arg2 or Predarg.

DRAFT DRAFT DRAFT DRAFT DRAFT DRAFT DRAFT DRAFT DRAFT DRAFT DRAFT DRAFT DRAFT DRAFT DRAFT DRAFT DRAFT DRAFT

Given these relationships in the content potential, we can define a **description** D_{E_0} of an entity-node E_0 as a sequence of entity-chains $C_1 \ldots C_n$ such that

- the focus of C_1 is E_0;
- each subsequent chain is linked via a resumption to an earlier chain;
- no chain begins more than p fact-nodes away from the fact-node which introduces it. (We currently set p to 8.)

We conclude by defining a **text** describing a complex entity E as a sequence of **descriptions** $D_{e1} \ldots D_{en}$, where $e1 \ldots en$ is an ordered set of subparts of E.

7 Algorithms and heuristics for content selection and structuring

In this section, we briefly outline our algorithms for selecting and structuring fact-nodes for a single description beginning from an initial entity-node E_0.

In the content selection process, a tree of fact-nodes is constructed whose root is a fact-node with E_0 as its Arg1, using all legitimate ways of moving from one fact-node to another. Heuristics about which nodes to include relate to meta-annotations on fact-nodes about their educational importance, their likely interest to the user, and the likelihood of their already being known, as well as to their distance from the original node.

Given the selected tree of fact-nodes, we extract from the text potential a complete subgraph containing *all* relation-nodes and *all* entity-nodes linking pairs of facts within this set. This forms the input to the text structuring algorithm. The subgraph contains the starting point for an approach like that of Marcu (1996); we have a set of all facts, and of all relations between pairs of facts. It differs from Marcu firstly in that the ELABORATION relation is not included in the set of relations; and secondly in having an explicit and distinct representation of all possible entity-based connections as well as of all possible relations.

The text structuring algorithm proceeds in a number of steps. We begin by dividing the fact-nodes into groups with the same Arg1, forming a set of initial entity-chains, and calculate

DRAFT DRAFT DRAFT DRAFT DRAFT DRAFT DRAFT DRAFT DRAFT DRAFT DRAFT DRAFT DRAFT DRAFT DRAFT DRAFT DRAFT

all possible resumption relations between these chains. We then perform an exhaustive search for the best rs-tree of depth n (currently set to 3) that can be formed with all the available fact-nodes, subject to the constraint that it can be added to one of the chains, or related to one by resumption, when the facts it uses up have been removed from the chains. (Heuristics for evaluating trees include preferences for some relations over others, for bushy trees, and for trees where one relation is not expanded as itself.) This process of adding the best possible rs-tree is iterated for the fact-nodes remaining in the chains, until no more legal trees can be added.

We then determine the ordering of entity-chains. If an entity-chain is short, we consider **NP-aggregation**, whereby it is incorporated as a relative clause whose head is the NP in the fact-node which introduces it. Otherwise, a chain is placed as close as possible to the chain which introduces it. If two chains are both resumptions from a single chain, we consider whether there is a fact-node F in one chain which could feature as joint focus return from a fact-node in the other; and if so, we move F to the end of the other chain and position it second. If C_1 is a resumption from C_0 but is not immediately adjacent to it, we also consider moving the fact-node which introduces C_1 to the beginning of C_1. Finally, when the order of fact-chains is determined, we decide the order of units within each chain.

The resulting text structures are then transformed into surface text. The mapping between text structure and surface text is now quite close. Rhetorical relations are realised with appropriate connectives. Local entity-based moves provide a source of information for generating appropriate referring expressions, in accordance with the centering model, and for decisions about clefting. Finally, resumption relations between entity-chains are used to determine potential paragraph breaks.

An example text conforming to the constraints in the above model, which the architecture generates, is given below.

(16) (1) This piece is a necklace. (2) It was designed by a jeweller called Jessie King. (3) It was designed in 1905. (4) It is made of silver and enamel.

(5) Jessie King was a famous designer. (6) She was Scottish, (7) but she worked in London. (8) It was in London that this piece was made.

DRAFT DRAFT DRAFT DRAFT DRAFT DRAFT DRAFT DRAFT DRAFT DRAFT DRAFT DRAFT DRAFT DRAFT DRAFT DRAFT DRAFT DRAFT

(9) Like the previous piece, (10) this piece is in the Arts-and-Crafts style. (11) Although the previous piece had a simple shape, (12) Arts-and-Crafts style jewels tend to be elaborate; (13) for instance, this piece has detailed florals.

There are three entity-chains in the text: (a) spans 1–4, (b) spans 5–8, and (c) spans 9–13. Within these chains there are a number of local rs-trees: spans 6–7 (topnuc 7), spans 9–10 (topnuc 10), and spans 11–13 (topnuc 12). Resumptions occur from (b) to (a), and from (c) to (a). A joint focus return occurs between spans 7 and 8. The resulting text, with its topic-clause cleft seems a good optimisation of focus and relation-based constraints; we are currently evaluating the output in a museum setting to see if this is indeed so.

8 Conclusion

This paper suggests that the relation of object-attribute ELABORATION has different discourse-structural properties from other RST relations, which motivate it being treated in a different way. We propose a distinction between proposition-based moves and entity-based moves; the two types of move are formally expressed as constraints on traversal of the content potential graph. On the basis of these constraints, we formulate the notion of a description, in which non-local dependencies are allowed for entity-based moves, and a treatment of clefts in discourse is provided. We propose algorithms and heuristics for content selection and structuring on this basis.

It remains to be seen how well this model of text structure extends to registers beyond the descriptive genre we are considering. This question can be considered by comparing the goal structure in our texts with that in other texts. The communicative goal in our texts is to describe a domain of inter-related entities. (The genre considered in Sibun (1992) is somewhat similar.) In such cases, it is perhaps better to think of a *set* of goals to describe individual entities, rather than a single overriding goal. The relationships between these goals are not exactly hierarchical, as they are in texts whose aims are principally argumentative or instructional; rather, the model we are investigating is one where executing one descriptive goal may provide an *opportunity* for the execution of another. To what extent the non-local dependencies in our genre of text are a function of its looser goal structure is a question we

DRAFT DRAFT DRAFT DRAFT DRAFT DRAFT DRAFT DRAFT DRAFT DRAFT DRAFT DRAFT DRAFT DRAFT DRAFT DRAFT DRAFT

hope to address in future work. The fullest answer to such a question undoubtedly requires not just the kind of computational work we have been pursuing, but also new empirical work on human discourse production.

Acknowledgements

The first, third and fourth authors are at the Division of Informatics at the University of Edinburgh. The second author is at the Department of Computer Science at the University of Otago. The first author is an Engineering and Physical Sciences Research Council (EPSRC) Advanced Fellow, and this work was supported through EPSRC grant GR/K53321. The support of the Economic and Social Research Council for HCRC is also gratefully acknowledged. Our thanks to our collaborators in the Education Department at the National Museums of Scotland, and to our referees for their helpful comments on earlier versions of this paper. Special thanks to the organisers and audience at the international workshop on text representation held in Utrecht in July 1997, where some of the ideas described here were first presented.

References

Ball, C. N. (1994). The origins of the informative-presupposition *it*-cleft. *Journal of Pragmatics*, **22**, 603–628.

Collins, P. (1991). *Cleft and Pseudo-Cleft Constructions in English*. Routledge, London and New York.

Delin, J. and Oberlander, J. (1995). Syntactic constraints on discourse structure: the case of *it*-clefts. *Linguistics*, **33**, 465–500.

Gazdar, G. (1979). *Pragmatics: Implicature, Presupposition, and Logical Form*. Academic Press, New York.

Grosz, B. J. and Sidner, C. L. (1986). Attention, intentions, and the structure of discourse. *Computational Linguistics*, pages 175–203.

DRAFT DRAFT DRAFT DRAFT DRAFT DRAFT DRAFT DRAFT DRAFT DRAFT DRAFT DRAFT DRAFT DRAFT DRAFT DRAFT DRAFT DRAFT

Grosz, B. J., Joshi, A. K., and Weinstein, S. (1995). Centering: A framework for modeling the local coherence of discourse. *Computational Linguistics*, **21**(2), 203–225.

Gundel, J. (1985). 'Shared knowledge' and topicality. *Journal of Pragmatics*, **9**, 83–97.

Gundel, J. (1988). Universals of topic-comment structure. In M. Hammond, E. A. Moravscik, and W. J. R., editors, *Studies in Syntactic Typology*. John Benjamins, Amsterdam and Philadelphia.

Halliday, M. A. K. and Hasan, R. (1976). *Cohesion in English*. Longman, London and New York.

Hedberg, N. (1990). *Discourse Pragmatics and Cleft Sentences in English*. Ph.D. thesis, University of Minnesota. UMI order number 9109340.

Heim, I. (1982). *The semantics of definite and indefinite noun phrases*. Ph.D. thesis, University of Massachusetts. Distributed by Graduate Linguistic Student Association.

Hobbs, J. R. (1985). On the coherence and structure of discourse. Technical Report CSLI-85-37, Center for the Study of Language and Information, Stanford University.

Kittredge, R., Korelsky, T., and Rambow, O. (1991). On the need for domain communication knowledge. *Computational Intelligence*, **7**, 305–314.

Mann, W. C. and Thompson, S. A. (1988). Rhetorical structure theory: A theory of text organization. *Text*, **8(3)**, 243–281. Also available as Tech Report RR-87-190, USC Information Sciences Institute, Marina del Rey, CA.

Marcu, D. (1996). Building up rhetorical structure trees. In *Proceedings of the AAAI annual meeting*.

Marcu, D. (1997). *The rhetorical parsing, summarisation and generation of natural language texts*. Ph.D. thesis, Department of Computer Science, University of Toronto.

McKeown, K. R. (1985). *Text Generation: Using Discourse Strategies and Focus Constraints to Generate Natural Language Text*. Cambridge University Press.

DRAFT DRAFT DRAFT DRAFT DRAFT DRAFT DRAFT DRAFT DRAFT DRAFT DRAFT DRAFT DRAFT DRAFT DRAFT DRAFT DRAFT

Mooney, D., Carberry, M. S., and McCoy, K. F. (1990). The generation of high-level structure for extended explanations. *COLING 90*, **2**, 276–281.

Moser, M., Moore, J., and Glendening, E. (1996). Instructions for coding explanations: Identifying segments, relations and minimal units. Unpublished manuscript.

Oberlander, J. and Delin, J. (1996). The function and intepretation of reverse wh-clefts in spoken discourse. *Language and Speech*, **39**, 185–227.

Prince, E. (1978). A comparison of *it*-clefts and wh-clefts in discourse. *Language*, **54**, 883–906.

Prince, E. (1992). The ZPG letter: subjects, definiteness, and information-status. In S. Thompson and W. Mann, editors, *Discourse Description: Diverse analyses of a fundraising text*, pages 295–325. John Benjamins, Amsterdam.

Sibun, P. (1992). Generating text without trees. *Computational Intelligence*, **8**(1), 102–122.

Sidner, C. L. (1981). Focusing for interpretation of pronouns. *Computational Linguistics*, **7**(4), 217–231.

Vallduví, E. (1992). *The informational component*. Garland, NY.

van Kuppevelt, J. (1995). Discourse structure, topicality and questioning. *Journal of Linguistics*, **31**, 109–147.

Wilkinson, J. and Pearce, N. (1996). *Harmony and Contrast: A journey through East Asian art*. National Museums of Scotland, Edinburgh.

DRAFT DRAFT DRAFT DRAFT DRAFT DRAFT DRAFT DRAFT DRAFT DRAFT DRAFT DRAFT DRAFT DRAFT DRAFT DRAFT DRAFT DRAFT