

# Cb or not Cb? Centering theory applied to NLG

Rodger Kibble  
Information Technology Research Institute  
University of Brighton  
Lewes Road  
Brighton BN2 4GJ  
U.K.

Rodger.Kibble@itri.brighton.ac.uk

## Abstract

Centering theory (CT) has been mostly discussed from the point of view of interpretation rather than generation, and research has tended to concentrate on problems of anaphora resolution. This paper examines how centering could fit into the generation task, separating out components of the theory which are concerned with *planning* and *lexical choice*. We argue that it is a mistake to define a total ordering on the transitions CONTINUE, RETAIN, SHIFT and that they are in fact epiphenomenal; a *partial* ordering emerges from the interaction between cohesion (maintaining the same center) and salience (realising the center as Subject). CT has generally been neglected by NLG practitioners, possibly because it appears to assume that the center is determined according to *feedback* from the surface grammar to text planning, but we argue that this is an artefactual problem which can be eliminated on an appropriate interpretation of the CT rules.

## 1 What is Centering?

Centering theory (CT) is a theory of discourse structure which models the interaction of *cohesion* and *salience* in the internal organisation of a text. The main assumptions of the theory as presented by (Grosz et al 1995 (GJW), Brennan et al 1987) are:

1. For each utterance in a discourse there is precisely one entity which is the centre of attention or *center*.
2. There is a preference for consecutive utterances within a discourse segment to keep the same entity as the center, and for the most salient entity "realised" in an utterance to be interpreted as the center of the following utterance.

3. The center is the entity which is most likely to be pronominalised.

These principles will be more precisely explicated in Sect. 2.

CT has proved attractive to NLP researchers because of the elegance and simplicity of the core proposals; it provides a framework for analysing text without having to make tough decisions about what a particular utterance is "about", since all notions are defined in purely structural terms.

Much research in CT has concentrated on interpretation, particularly reference resolution, developing algorithms to resolve anaphoric expressions based on the assumption that the text is constructed according to Rules 1 and 2. So researchers have focussed on filling in details of the theory which were left unspecified: what counts as an utterance, and how should transitions be handled in complex sentences (Kameyama 1998; cf Suri and McCoy 1994)? how is salience ranking determined (Gordon et al 1993; Stevenson et al 1994; Strube and Hahn 1996)? what counts as "realisation" - does this include bridging references (Strube and Hahn op cit.)? how do centering transitions relate to discourse segment boundaries (Walker 1998, Passoneau 1998)?

In fact I will leave many of these issues aside for the purposes of this paper; I will not examine the empirical adequacy of CT, for which the reader is referred to papers cited above and others collected in Walker et al (1998). I will take a different approach, which is to examine how the elements of CT can be applied to the *planning* of texts, with the rules and constraints interpreted as instructions to a generator rather than a guide for interpretation. To avoid introducing too many complications I shall assume

	$Cb(U_n) = Cb(U_{n-1})$ or $Cb(U_{n-1})$ undefined	$Cb(U_n) \neq Cb(U_{n-1})$
$Cb(U_n) = Cp(U_n)$	Continue	Smooth Shift
$Cb(U_n) \neq Cp(U_n)$	Retain	Rough Shift

Figure 1: Centering Transitions

<b>Constraints</b>
C1. There is precisely one $Cb$ .
C2. Every element of $Cf(U_n)$ must be realised in $U_n$ .
C3. $Cb(U_n)$ is the highest-ranked element of $Cf(U_{n-1})$ that is realised in $U_n$ .
<b>Rules</b>
R1. If some element of $Cf(U_{n-1})$ is realised as a pronoun in $U_n$ then so is $Cb(U_n)$ . (Strong version: if $Cb(U_n) = Cb(U_{n-1})$ , a pronoun should be used.)
R2. <i>Continue</i> is preferred over <i>Retain</i> which is preferred over <i>Smooth Shift</i> which is preferred over <i>Rough Shift</i> .

Figure 2: Centering Constraints and Rules

a "canonical" formulation of CT as outlined by Walker et al. (1998, Chapter 1) and the schematic "consensus" generation architecture described by Reiter and Dale (Reiter 1994; Reiter and Dale 1997). This consists a "pipeline" of distinct tasks:

**Text Planning** - deciding the content of a message, and organising the component propositions into a text tree;

**Sentence Planning** - aggregating propositions into clausal units and choosing lexical items corresponding to concepts in the knowledge base, including referring expressions (RE);

**Linguistic realisation** which takes care of surface details such as agreement, orthography etc.

Previous researchers have implemented pronominalisation decisions using CT, and so have located Centering as part of RE generation (e.g., Dale 1992, Passoneau 1998), while Mittal et al (1998) have a "centering module" which forms part of Sentence Planning and seeks to realise the center as Subject in successive sentences. In what follows I will try to separate out the tasks which make up centering theory and argue that the way to implement CT is not as a discrete module but as a series of constraints on the various levels of the generation process from Text Planning

to RE generation. I shall also briefly note points of comparison with systems discussed by Cahill and Reape (1998) in a survey of applied NLG systems, and conclude with some remarks on the applicability of my proposals to the "reference architecture" envisaged by Cahill et al. (1999), RAGS (1999).

## 2 Transition rules

The main claims of CT are formalised in terms of  $Cb$ , the "backward-looking center",  $Cf$ , a list of "forward-looking centers" for each utterance  $U_n$ , and  $Cp$  or "preferred center", the most salient candidate for subsequent utterances.  $Cf(U_n)$  is a partial ordering on the entities mentioned (or "realised") in  $U_n$ , ranked by grammatical role, e.g. SUBJ > DIR-OBJ > INDIR-OBJ > COMP(s) > ADJUNCT(s).  $Cp(U_n)$  is the highest ranked member of  $Cf(U_n)$  (usually SUBJ), and is predicted to be  $Cb(U_{n+1})$ .  $Cf$  is partially ordered in most accounts, which leaves open the possibility that there is no unique  $Cb$ . Also, if two successive utterances have no referent in common the second will have no  $Cb$ .

Successive pairs of utterances are characterised in terms of *transitions* as defined in Figure 1; for instance if two consecutive utterances have the same  $Cb$ , and the  $Cb$  in the second utterance occurs in Subject position, this is classified as a CONTINUE transition. A text is judged to be coherent to the extent that transitions follow the preference ordering given in Rule 2 (Fig

2); and on the assumption that the text is coherent, pronominalisation is predicted to conform to Rule 1.

The notion of "realisation" is subdivided into "direct" and "indirect": an entity is directly realised in an utterance if it is the denotation of an overt NP (or a zero pronoun where this is syntactically licensed), while "indirect realisation" covers for example subsectional anaphora, possessive pronouns and inferential links such as bridging reference. Corpus-based investigations of CT have tended to concentrate on direct realisation, since annotation of indirect anaphoric links depends on theoretically-based decisions and it may be difficult to achieve reliability in this area. This has obvious implications for the resulting measure of coherence, which I return to below.

## 2.1 Salience and cohesion

Transitions are defined in terms of tests which I shall call cohesion:  $Cb(U_n) = Cb(U_{n-1})$ , and salience:  $Cb(U_n) = Cp(U_n)$ . There are four possible combinations, which are displayed in Fig. 1. The most preferred case is where both apply, namely CONTINUE, and the least preferred is where neither apply, ROUGH SHIFT. For the intermediate cases there are three logical possibilities: prefer cohesion (RETAIN), prefer salience (SMOOTH SHIFT) or allow both equally. There is no obvious way to settle this *a priori*, but Walker et al (1998, Ch. 1) stipulate that RETAIN is preferred over SMOOTH SHIFT. Evidence for this position is not conclusive, and indeed di Eugenio (1998), Passoneau (1998) and Hurewitz (1998) all found a higher percentage of shifts than retains. This suggests either that salience is a stronger requirement than cohesion or that it is easier to satisfy. That is, the languages studied (English and Italian) may be sufficiently flexible that there is usually some way to realise  $Cb$  as Subj (or first-mention) but on the other hand the same  $Cb$  can only be maintained for a finite number of utterances.

I suggest that these results should be treated with some caution since it is not clear that the authors have the same assumptions about the claims of CT or that what they are testing directly reflects formulations of CT in the more theoretical literature. For instance Passoneau (1998) refers to two variants of CT: "Version A" based on Brennan et al (1987) and "Version

B" taken from Kameyama et al (1993). Passoneau does not address the issue of direct vs indirect realisation and it appears from the examples given that she only takes account of entities realised by a full NP or (possibly null) pronoun. The analysis according to Version B results in a count of 52% NULL transitions, i.e. no  $Cb$ , which gives the impression that CT is in fact a rather poor measure of coherence. It is probable that a higher measure might have been obtained if Passoneau had allowed entities to be added to the  $Cf$ 's by inference, as discussed in (Brennan et al, op cit.). It is of course impossible to verify this without access to the original texts, but it is instructive to consider the following example from Strube and Hahn (1996):

- (1) a. Ein Reserve-Batteriepack versorgt den 316LT ca. 2 Minuten mit Strom.  
(A reserve battery pack - supplies - the 316LT - ca. 2 minutes - with power.)
- b. Der Status des Akkus wird dem Anwender angezeigt.  
(The status of - the accumulator - is - to the user - indicated.)

S & H treat *Akku* in the (b) sentence as indirectly realising the 316LT (a kind of computer) in the (a) sentence, so the latter becomes the  $Cb$  of (b) resulting in a CONTINUE transition. If the authors had only taken account of direct realisations this would be analysed as a NULL transition.

Furthermore, a preponderance of shifts over continues may reflect the domain and content rather than the organisation of a text. In fact it can be seen that sequences of smooth shifts are rather natural in certain kinds of narrative or descriptive texts; see example (2).

- (2) The name of your medicine is Rhinocort Aqua.  
It contains budesonide.  
This is one of a group of medicines called corticosteroids. SMOOTH SHIFT  
These can help to relieve the symptoms of hay fever or rhinitis. SMOOTH SHIFT  
  
(pharmaceutical leaflet)

This does not appear to be an incoherent text,

but there is no way that the content could be rearranged to turn the shifts into continues.

## 2.2 Deconstructing the transitions

Strube and Hahn (1996) question the canonical ordering of transitions, partly on the grounds that this ordering fails to predict the RETAIN - SHIFT pattern which has been claimed by many researchers to signal a segment boundary or the introduction of a new "discourse topic". (See Section 3.2 below.) Recall that the  $C_p$  is defined as the most salient entity realised in  $U_n$ , which is predicted to be the  $C_b$  of  $U_{n+1}$ . However this "prediction" is not in fact cashed out in the rules for centering transitions, which take no account of whether  $C_p(U_n)$  is actually realised in  $U_{n+1}$ . S & H (op cit., p. 275) propose the principle of cheapness which is satisfied if  $C_p(U_n) = C_b(U_{n+1})$ . "Cheapness" is claimed to minimise the inferential costs of processing sequences of utterances, and is proposed as a constraint on pairs of successive transitions as a replacement for the canonical orderings in Rule 2.

We call a transition pair *cheap* if the backward-looking center of the current utterance is correctly predicted by the preferred center of the immediately preceding utterance, i.e.,  $C_b(U_i) = C_p(U_{i-1})$ ...

In fact it turns out that although cheapness appears to be a sensible principle, it does not neatly partition the types of transition pairs; in particular, this principle does not necessarily hold of all RETAIN - SMOOTH SHIFT sequences. S & H propose to rectify this by redefining the transitions, with an additional test  $C_b(U_i) = C_p(U_{i-1})$  to subdivide CONTINUE and SMOOTH SHIFT (Strube, p.c.; Strube and Hahn forthcoming).

In what follows I will also argue against the canonical ordering though on different grounds: one cannot in general predict a preference for Retain over Shift, for the simple reason that there is no point at which the choice between these two alternatives arises. Rather, at different points in the generation process there is a choice between maintaining the same  $C_b$  or choosing a new one, and a choice of which entity ( $C_b$  or non- $C_b$ ) to make salient. So the various transition types emerge in a partial ordering

from the interaction between salience and cohesion. Note that if we also include Strube and Hahn's "cheapness" principle, there is potential competition with salience in cases where  $C_b(U_n) \neq C_b(U_{n+1})$ . That is, we will need a way to decide which entity to realise as  $C_p$  in cases where there are two candidates, one of which is the current  $C_b$  and the other the  $C_b$  of the following sentence. In the remainder of this paper I will not directly incorporate the "cheapness" principle but will suggest that similar results are obtained with an appropriate interpretation of Constraint 3 in the context of generation.

## 3 Architecture

If we decompose the rules and constraints into separate specifications we see that they potentially fall under quite different headings in the schematic architecture described above. The transitions mentioned in Rule 2 are defined in terms of two principles which I have called cohesion (maintaining the same  $C_b$  from one utterance to the next) and salience (realising the  $C_b$  as  $C_p$ , normally Subject). If we consider these as planning operations, cohesion naturally comes under Text Planning and salience under Sentence Planning, while Rule 1 concerning pronominalisation falls under the Referring Expression component of Sentence Planning.

### 3.1 Text Planning

A text planner which operated according to CT would seek to order clauses within a segment to maintain the same  $C_b$  in a sequence of clauses. There are two related issues which complicate this project: firstly, Constraint 3 in Fig. 2 implies a requirement for *feedback* to determine the  $C_b$ . In addition, there is a potential conflict between top-down, hierarchical RST-type planning and sequential centering rules.

#### 3.1.1 Identifying the $C_b$

Constraint 3 states that "for each utterance  $U_i$  in a discourse segment  $D$ ...

The center,  $C_b(U_i, D)$  is the highest-ranked element of  $C_f(U_{i-1}, D)$  that is realised in  $U_i$ ." (Walker et al 1998:3).

There are two different implementation strategies which could satisfy this constraint:

1. Take the ranking of  $Cf(U_{i-1}, D)$  as given and use this to identify the  $Cb$ .

2. Take the  $Cb$  as given and plan the realisation of  $U_{i-1}$  to make this entity the highest-ranked.

The first strategy is clearly appropriate for interpretation (cf Brennan et al 1987) but for generation the issue is less clear-cut. Either the generator "interprets" its own output to designate  $Cb$  in terms of the grammatical structure of the previous utterance, in which case there have to be separate principles for deciding on the grammatical structure, or  $Cb$  is independently defined in the text plan and this information is used to plan the sentence structure.

According to the pipelining principle information cannot flow 'backwards' between tasks. In a pipelined system, the realisation of an entity as  $Cp$  may have the effect of setting up an expectation on the reader's part that this entity will be  $Cb$  of the following utterance, but it cannot influence the decision made by text planning. This would mean that the  $Cb$  will no longer be defined as in Constraint 3 but must be independently designated by the text planner as the centre of attention in an utterance. In fact the resulting distribution of tasks would be rather similar to the Gossip system (Carcagno and Iordanskaja 1989) as described by Cahill and Reape (1998):

First, the planner produces "sentence-sized semantic nets" which it marks with theme-rheme information. ... Furthermore, the theme/rheme constraints influence clause ordering, ... pronominalisation ... and lexical choice.

An implementation of CT with feedback seems likely to fit more naturally into an *incremental* architecture, where generation tasks may be carried out on an utterance-by-utterance basis in contrast to top-down generation of a text tree for an entire discourse. In incremental systems it is possible in principle to plan the content of an utterance according to which entity is currently made salient by its surface grammatical role, whereas this would not in general be possible in a top-down pipelined system. In fact it turns out that incremental generators tend to perform sentence planning

incrementally but not text planning. (See e.g. Reithinger 1991, DeSmedt and Kempen 1991.)

In fact I would argue that the feedback problem is only an artefact of an interpretation of Constraint 3 as an implication rather than a constraint. The "implicational" interpretation is that if an entity  $a$  is  $Cp$  of  $U_n$  and is realised in  $U_{n+1}$ , it should be designated as  $Cb$  of  $U_{n+1}$ . The declarative interpretation is non-directional and simply equates  $Cb$  of  $U_{n+1}$  with the most salient entity realised in  $U_n$  which is also realised in  $U_{n+1}$ . The way to implement this while keeping to the pipelining principle is to assume that the text planner independently designates the "theme", "topic" or intended centre of attention in each clause, which is marked as  $Cb$  if it is realised in the previous clause, and to have the Sentence Planner promote an entity to salience in  $U_n$  if it is  $Cb$  of  $U_{n+1}$ . So the text planner should annotate each clausal node  $U_n$  in the text tree with the following information:

$Cb$  of  $U_n$   
 $Cb$  of  $U_{n-1}$   
 $Cb$  of  $U_{n+1}$

The sentence planner will then make use of this information to decide on pronominalisation according to the values of the current and previous  $Cb$ , and on promotion of arguments to salience depending on the current and *following*  $Cb$ . Some concrete proposals are discussed in Section 3.2, "Sentence Planning". The general division of labour is outlined in Fig. 3.

### 3.1.2 Top-down Text Planning

The reader may be disappointed that no independent definition of "topic" or "theme" is offered. In fact we consider that this may be outside the scope of CT. When used for interpretation, CT offers a set of rules of thumb to guide the system in identifying the centre of attention in each utterance and finding probable antecedents for anaphors, but the notion of "centerhood" is not defined separately from these rules. When used for generation, the most CT can offer is to take the topic or theme as *given* and exploit the centering rules and constraints to construct the text in a way which foregrounds these entities and enables the user to correctly identify antecedents. So CT has to sit on top of an independently specified treatment of information structure. One candidate is Strube and Hahn's (1996) reformulation of the notions of

<b>Text Planning</b>
<ol style="list-style-type: none"> <li>1. <i>Content determination.</i></li> <li>2. <b>Discourse planning:</b> order clauses <math>U_i</math> (within segments) to maximise continuity of reference.</li> </ol> <p>For each clause <math>U_n</math>:</p> <ul style="list-style-type: none"> <li>o Designate at most one argument as <math>Cb(U_n)</math>, which must be an argument of <math>U_{n-1}</math>. (If intersection of <math>Cfs(U_n)</math> and <math>Cfs(U_{n-1})</math> has only one member, that member is <math>Cb</math>.)</li> <li>o Annotate clause node with IDs of <math>Cb(U_n)</math>, <math>Cb(U_{n-1})</math> and <math>Cb(U_{n+1})</math>.</li> </ul>
<b>Sentence Planning</b>
<ol style="list-style-type: none"> <li>1. <i>Sentence aggregation.</i></li> <li>2. <b>Lexicalisation:</b> select verb form for <math>U_n</math> so that <ol style="list-style-type: none"> <li>a. <math>Cb(U_{n+1})</math> is most grammatically salient of intersection of <math>Cfs(U_n)</math> and <math>Cfs(U_{n+1})</math>;</li> <li>b. subject to (a), <math>Cb(U_n)</math> is realised in most salient available position.</li> </ol> </li> <li>3. <b>Referring expression generation:</b> working hypotheses <ul style="list-style-type: none"> <li><math>Cb(U_n)</math> may be pronominalised if: <ul style="list-style-type: none"> <li>o <math>Cb(U_n) = Cb(U_{n-1})</math> (GJW 83)</li> <li>o <math>Cb(U_n) = Cp(U_{n-1})</math> (Brennan 98)</li> </ul> </li> </ul> </li> </ol>

Figure 3: Locating centering tasks in the pipeline

*theme* and *rheme*. Another possibility, which is currently under investigation, is to experiment with the effects of rhetorical structure on choice of  $Cb$ .

I've assumed above that the task of maintaining continuity of reference can be located in the Text Planner. That is, the TP would be responsible both for annotating the  $Cb$  in each utterance, and for organising the text so that the same  $Cb$  is maintained over a sequence of clauses. However, according to Reiter and Dale (1997) a more common method of structuring text is to make use of "discourse relations" such as those described by Mann and Thompson (1987), which do not explicitly take continuity of reference into account. Richard Power (p.c.) has proposed that the implementation of CT in an RST-based text planner can be treated as an optimisation problem. That is, the text plan is initially taken to be a tree structure with discourse relations defined on adjacent nodes but at most a partial specification of linear order. The problem will then consist of selecting a  $Cb$  for each propositional leaf node in such a way as to maximise the coherence of the text according to centering rules. This is an area of active research. Cheng (MS) has proposed a similar strategy for maintaining local coherence in a text planner using a genetic algorithm.

Another issue is whether the CT rules, which assume a "flat" sequence of utterances, will re-

main valid for a hierarchically structured text plan. In fact it is an open research question whether CT should operate in this manner or whether the rules should be reformulated to take account of dominance in addition to precedence; different positions are taken by Kameyama (1988) and Suri and McCoy (1994).

### 3.2 Sentence planning

According to the re-interpretation of CT which was sketched above, Sentence Planning may promote an entity for salience if it is the  $Cb$  of the current or following utterance. There is clearly potential competition between these two factors which will be discussed shortly. The preference for realising  $Cb$  as  $Cp$  can be implemented by choosing a verb form which projects  $Cb$  in subject position. Some possibilities are transfer alternation: *buy/sell, give/receive, borrow/lend* etc, or passivisation: *your doctor may prescribe this medicine for gout vs this medicine may be prescribed for gout*.

If we compare the attested example (2) with the constructed (3) it is clear that the former reads more naturally:

- (3) Hypoglycaemia ( $Cb$ ) may cause faintness, sweating, shaking, weakness and confusion. It may be due to lack of food or too high a dose of the medicine. CONTINUE  
It can be put right by eating or drinking something sugary. CONTINUE

(pharmaceutical leaflet)

- (4) ... Hypoglycaemia may cause faintness, sweating, shaking, weakness and confusion. A lack of food or too high a dose of the medicine may cause it. RETAIN  
Eating or drinking something sugary can put it right. RETAIN  
(modified example)

Hurewitz (1998) examined the use of passives to promote cohesion and salience and found that in both written texts and speech approximately 75% of passives had either the CONTINUE or SMOOTH-SHIFT transition. In each case the effect is to promote *Cb* to Subject in accordance with the salience principle. (For written texts this proportion was not significantly different from a control sample, whereas with the spoken passages the proportion was slightly higher than in the control.)

One system which explicitly makes use of CT is the Caption Generation System (CGS) reported in (Mittal et al 1998). This system has a separate "centering module" which orders arguments within a clause to improve coherence of a text but does not influence the order of clauses. Thus only the salience principle is implemented and the centering task is located as part of Sentence Planning: the specialised Centering Module receives control after clauses have been ordered and aggregated (op cit:454). The strategy adopted is to keep "the highest-ranking forward-looking center of the first clause of the segment ... as the  $Cp(U_i)$  of all the following clauses in the same segment" (op cit:456; my emphasis). Clearly this strategy is unlikely to generalise to a variety of domains.

As mentioned above there is a potential conflict between Constraint 3 (make  $Cb(U_{n+1})$  salient) and salience (make  $Cb(U_n)$  salient). As noted in Sect. 2.2 there may also be competition between salience and Strube & Hahn's cheapness principle, which can be seen as a stronger version of C3. There are different ways this conflict could be tackled in the cases where it arises and I will consider one of them, which is to let C3 win out over salience.

Consider a text with four clauses  $U_1 - U_4$  which all have *a* and *b* among their arguments. Let *a* be the *Cb* of  $U_1 - U_2$  and *b* the *Cb* of  $U_3 - U_4$ . According to salience and C3, *b* will be  $Cp$  of

$U_3$  since it is *Cb* of that clause and the following one. For  $U_2$  there is competition between *a* and *b* to be  $Cp$ , and this is decided by C3 in favour of *b*. Finally<sup>1</sup> *a* is chosen as  $Cp$  of  $U_1$ . The result is as follows:

$U_1 : Cb = a, Cp = a$

$U_2 : Cb = a, Cp = b$

$U_3 : Cb = b, Cp = b$

$U_4 : Cb = b, Cp = b$

In terms of the conventional transitions this works out as

$U_1/U_2$ : RETAIN

$U_2/U_3$ : SMOOTH SHIFT

$U_3/U_4$ : CONTINUE

This is consistent with Strube and Hahn's (1996) observation that "a RETAIN transition ideally predicts a SMOOTH SHIFT in the following utterance". Brennan et al (1987) make a very similar claim:

A computational system for *generation* would try to plan a retention as a signal of an impending shift, so that after a retention, a shift would be preferred rather than a continuation.

Grosz et al (1995) give the following example of the RETAIN - SHIFT pattern:

- (5) a. John has had trouble arranging his vacation.  
b. He (*Cb*; John) cannot find anyone to take over his responsibilities.  
c. He (*Cb*; John) called up Mike yesterday to work out a plan. CONTINUE  
d. Mike has annoyed him (*Cb*; John) a lot recently. RETAIN  
e. He (*Cb*; Mike) called John at 5 am on Friday last week. SHIFT

Under the approach outlined here, which assumes that the *Cb* is independently designated, the system does not need to plan particular transition types or even to know about them; the desired effects come about as a result of local decisions by the sentence planner using information from the text planner.

Example (5) incidentally illustrates a limitation of CT in its canonical version: the theory

<sup>1</sup>The word "finally" does not imply that these decisions are taken in sequence.

correctly predicts the pronominal choices in (5c - e) but has nothing to say about the decision to make *Mike* rather than *John* the Subject of (5d). In fact, we can construct a variant of this text which follows centering rules more faithfully, though it does not read any more naturally:

- (6) a. John has had trouble arranging his vacation.
- b. He cannot find anyone to take over his responsibilities. (Cb = John)
- c. He called up Mike yesterday to work out a plan.  
(CONTINUE; Cb = John)
- d. He has been pretty annoyed with Mike recently.  
(CONTINUE; Cb = John)
- e. He got a call from him (Mike) at 5 am on Friday last week. (CONTINUE; Cb = John)

If we examine the discourse structure of example (5), it seems that the discourse as a whole is about *John* but (5d,e) form a parenthetical section which tells us something about *Mike*. So although a blind application of centering rules would judge (6) to be more coherent, (5) is in fact maximally coherent within the constraints of the structure of the discourse.

To summarise: we assume that for each utterance  $U_n$  the text planner has identified  $Cb(U_n)$ ,  $Cb(U_{n-1})$  and  $Cb(U_{n+1})$ . The task for the sentence planner is to select a verb form or some other syntactic device to realise  $Cb(U_{n+1})$  as the most salient of those entities which are realised in  $U_n$  and  $U_{n+1}$ , and subject to this, to realise  $Cb(U_n)$  in the most salient position available. (See Figure 3.) So for example if  $Cb(U_n)$  is *not* to be realised in  $U_{n+1}$ , it will normally be realised as  $Cp(U_n)$ ; but if it *is* to be realised in  $U_{n+1}$  then the  $Cb$  of  $U_{n+1}$  will be at least as highly ranked in  $U_n$  as the  $Cb$  of  $U_n$ . This strategy predicts the realisation of *Mike* rather than *John* as  $Cp$  in (5d) above.

### 3.3 Referring Expressions

The contribution of CT to Referring Expression (RE) generation is to decide on pronominalisation. Rule 1 (Fig. 2) which concerns pronom-

inalisation has a strong and a weak formulation: the strong one is that a pronoun should be used for the  $Cb$  if it is the same as the  $Cb$  of the previous utterance, the weak one is that the  $Cb$  must be pronominalised if anything is. In the context of generation it is probably safer to use the strong version. Brennan (1998) proposes, arguing from corpus analysis, that the  $Cb$  should be pronominalised only if it is  $Cp$  of the previous utterance. Robert Dale's EPICURE system employs the terminology of CT in connection with RE generation, as does the ILEX system reported in (O'Donnell et al 1998). Both these systems implement a variant of Rule 1 to determine whether to pronominalise the *center* (Dale) or  $Cb$  (ILEX), though in neither case is the center identified according to the standard apparatus of CT. In ILEX the  $Cb$  is designated by the text planner without reference to the content of the previous sentence, and it may be pronominalised if it is the same as the previous sentence's "Cb". Dale's EPICURE identifies the center with "that entity which is the result of the previously described operation" (Dale 1992:170). Passoneau (1998) constructed input for a prototype generator by "hypothesising" a  $Cb$  for each proposition in a text based on the salience of entities in a situation. Passoneau's system uses centering constraints to decide whether to realise entities as definite pronouns, minimal NPs or full NPs.

## 4 Conclusions and future work

CT has developed primarily as a tool for analysing the structure of a given text and identifying the most likely candidates for anaphora resolution. In this paper we sought to determine whether the principles underlying the constraints and rules of the theory can be "turned round" and used as planning operators for generating coherent text. As a side-effect of this enterprise we have articulated a "streamlined" formulation of CT in terms of the principles of salience and cohesion, and argued that the preferences for the different transition types emerge in a partial ordering from the interaction between these principles. These principles are rather heterogeneous, a fact which is obscured by combining them in the transition definitions, and can be implemented as encapsulated tasks distributed between text planning,



sentence planning and RE generation. It may turn out that individual components such as the cohesion principle do not need to be explicitly stated but emerge as by-products from higher-level text planning.

As noted at various points in this paper CT has never been more than partially implemented in NLG systems. This may be due to a belief on the part of NLG practitioners that CT gets things the wrong way round, by relying on surface grammatical realisation to determine the centre of attention in an utterance. If this belief is commonly held (and anecdotal evidence suggests that it is), I argue that it is mistaken. In interpretation systems the principles of CT guide the system in identifying the centre of attention and in choosing likely referents for anaphors. In NLG systems, if there is a notion of "topic" or "theme" this should be designated by the text planner, while the CT rules allow the sentence planner to promote this entity to salience to keep it as the user's "centre of attention".

However, a more fundamental explanation for the neglect of CT in the generation literature is provided by the fact that a faithful implementation in a pipelined system turns out to require an independent way of designating the central entity in a proposition, and this itself is a problem which has not had much attention in the development of NLG systems<sup>2</sup>. So the next stage in this research will concentrate on developing a characterisation of *Cb* based on semantic content and information structure, taking account of e.g. the proposals of Strube and Hahn (1996) and experimenting with optimisation algorithms as discussed in section 3.1.2 above.

As mentioned above the Reiter model has been questioned by members of the RAGS project in Brighton and Edinburgh who are actively engaged in developing a "reference" architecture for NLG. (See Cahill et al., 1999, RAGS 1999.) To date the group has concentrated on specifying the data structures which are required at various stages of the generation task and has identified a number of discrete functions such as rhetorical structuring, aggregation,

<sup>2</sup>This applies at least to *applied* systems; see Cahill and Reape 1998. One exception appears to be the GOS-SIP system described in Caragno and Iordanskaja 1989.

coherence etc., without specifying a strict order for the execution of these functions. It is too early to assess how the proposals of this paper would fit into the RAGS scheme, but I anticipate that, as with the Reiter architecture, the conclusion would be that referential coherence is not the task of a discrete module but imposes constraints on a number of different modules.

It has been noted that the way the "realise" relation is interpreted can have significant implications for the coherence of a text as measured by CT, and that corpus analysis has often concentrated on directly realised *Cb*'s. An exception is Hahn, Strube and Markert's (1996) treatment of bridging reference, or "textual ellipsis" in their terminology. As these authors note there have been rather few implemented systems which are able to interpret bridging references in a principled way, and research in NLG is particularly weak in this area. So, a faithful implementation of CT in generation systems will depend in part on progress in the generation of bridging references. This is an area for future research.

It is intended that the procedures described in this paper will be implemented in ICONOCLAST, an authoring tool which enables domain experts to create a knowledge base through a sequence of interactive choices and generates hierarchically structured text according to various stylistic constraints (See Power and Scott 1998).

### Acknowledgements

This work was carried out as part of the GNOME project (Generating Nominal Expressions) which is a collaboration between ITRI in the University of Brighton and the HCRC in the universities of Edinburgh and Durham, funded by the EPSRC under grant reference GR/L51126. I would like to thank ITRI and GNOME colleagues for helpful feedback, particularly Christy Doran, Renate Henschel, Richard Power and Kees van Deemter, as well as two anonymous referees.

### References

- S Brennan, M Walker Friedman and C Polard 1987. A Centering Approach to Pronouns. In *Proc. 25th ACL* :115-62.
- L Cahill, C Doran, R Evans, C Mellish, D Paiva, M Reape and D Scott 1999. In search of a reference architecture for NLG systems. In *Proc.*

*EWNLG'99.*

- L Cahill and M Reape 1998. Component tasks in Applied Natural Language Generation Systems. RAGS Project Deliverable, available at [www.itri.brighton.ac.uk/projects/rags](http://www.itri.brighton.ac.uk/projects/rags).
- D Caragno and L Iordanskaja 1989. Content Determination and Text Structuring in GOSSIP. In *Extended Abstracts of ENLG'89*.
- H Cheng MS. Experimenting with the Interaction between Aggregation and Text Planning. Unpublished paper, Division of Informatics, University of Edinburgh.
- R Dale 1992, *Generating Referring Expressions*, Cambridge, MA/London:MIT Press.
- K De Smedt and G Kempen 1991. Segment Grammar: a Formalism for Incremental Sentence Generation. In C Paris, W Swartout and W Mann (eds), *Natural Language Generation in Artificial Intelligence and Computational Linguistics*.
- B Di Eugenio 1996. The discourse functions of Italian subjects: a centering approach. In *Proc. COLING96*.
- B Di Eugenio 1998. Centering in Italian. In Walker et al (eds) *Centering Theory in Discourse*.
- P Gordon, B Grosz and L Gilliom 1993. Pronouns, Names and the Centering of Attention in Discourse. *Cognitive Science*, 17/3:311-47.
- B Grosz, A Joshi and S Weinstein 1995. Centering: a framework for modelling the local coherence of discourse. *Computational Linguistics*, 21/2: 203-25.
- U Hahn, M Strube and M Markert 1996. Bridging Textual Ellipses. In *Proc. of COLING-96*.
- F Hurewitz 1998. A Quantitative Look at Discourse Coherence. In Walker et al (eds) *Centering Theory in Discourse*.
- M Kameyama 1998. Intrasentential Centering: A Case Study. In Walker et al (eds) *Centering Theory in Discourse*.
- M Kameyama, R Passoneau and M Poesio 1993. Temporal Centering. In *Proc. of 31st ACL*.
- W Mann and S Thompson 1987. Rhetorical Structure Theory: A Theory of Text Organisation. In L Polanyi (ed.), *The Structure of Discourse*.
- V Mittal, J Moore, G Carenini and S Roth 1998. Describing Complex Charts in Natural Language: A Caption Generation System. *Computational Linguistics*, 24/3:431-468.
- R Power and D Scott 1998. Multilingual authoring using feedback texts. In *Proc. COLING/ACL*

'98

- RAGS 1999. The RAGS Project. Towards a Reference Architecture for Natural Language Generation Systems. Technical report ITRI-99-14, ITRI, University of Brighton. Available at [www.itri.brighton.ac.uk/projects/rags](http://www.itri.brighton.ac.uk/projects/rags).
- E Reiter 1994. Has a consensus NL generation architecture appeared, and is it psycholinguistically plausible? In *Proc. INLG 7:163-70*.
- E. Reiter and R. Dale (1997). Building Applied Natural-Language Generation Systems. *Journal of Natural-Language Engineering* 3:57-87.
- N Reithinger 1991. POPEL - A Parallel and Incremental Natural Language Generation System. In C Paris, W Swartout and W Mann (eds), *Natural Language Generation in Artificial Intelligence and Computational Linguistics*.
- R Stevenson, R Crawley and D Kleinman 1994. Thematic roles, focus and the representation of events. *Language and Cognitive Processes*, 9:519-48.
- M Strube and U Hahn 1996. Functional Centering. In *Proceedings ACL* 34:270-77.
- M Strube and U Hahn 1999. Functional Centering: Grounding Referential Coherence in Information Structure. To appear in *Computational Linguistics*.
- L Suri and K McCoy 1994. RAFT/RAPR and Centering: A Comparison and Discussion of Problems Related to Processing Complex Sentences. *Computational Linguistics* 20/2:301-17.
- M Walker 1998. Centering, Anaphora Resolution and Discourse Structure. In Walker et al (eds) *Centering Theory in Discourse*.
- M Walker, AK Joshi and E Prince (eds) 1998. *Centering Theory in Discourse*. Oxford: Clarendon Press.