Taking Time to Structure Discourse: Pronoun Generation Beyond Accessibility

Kathleen F. McCoy
Dept. of Computer and Information Sciences
University of Delaware
103 Smith Hall
Newark, DE 19716, USA
mccoy@cis.udel.edu

Michael Strube
Institute for Research in Cognitive Science
University of Pennsylvania
3401 Walnut Street, Suite 400A
Philadelphia, PA 19104, USA
strube@linc.cis.upenn.edu

Abstract

In order to produce coherent text, natural language generation systems must have the ability to generate pronouns in the appropriate places. In the past, pronoun usage was primarily investigated with respect to the accessibility of referents. That is, it was assumed that a pronoun should be generated whenever the referent was sufficiently accessible so as to make its resolution easy. We found that such an explanation does not seem to account well for the patterns of pronoun usage found in naturally occurring texts. We present an algorithm for generating appropriate anaphoric expressions which takes into account the temporal structure of texts (as a discourse structuring device) and knowledge about ambiguous contexts. Other important factors in our algorithm are sentence boundaries and the distance from the last mention of the anaphor. We back up our hypotheses with some empirical results indicating that our algorithm chooses the right referring expression in 85% of the cases.

Introduction

Anaphoric expressions are an important component to generating coherent discourses. While there has been some work on generating appropriate referring expressions, little attention has been given to the problem of when a pronoun should be used to refer to an object. In most instances the assumption has been that a pronoun should be generated when referring to a discourse entity that is highly prominent (accessible). However, a study of naturally occurring texts reveals that factors beyond accessibility must be taken into account in order to explain the patterns of pronoun use found.

In this paper we attempt to investigate factors that might influence the use of pronominal reference forms. In particular, we attempt to answer the question of when it is appropriate to generate a pronoun versus some other kind of definite description (i.e., a definite noun phrase or proper name) when realizing an anaphoric expression. Pronouns are prevalent in text and play a role in text coherence, yet pronoun generation has gotten very little attention. To date most of the research involving pronouns in text has concentrated on the problem of pronoun resolution (needed for natural language understanding). While it is likely that the work on pronoun resolution may be relevant to the problem of pronoun generation (one would not want to generate a pronoun that the reader could not resolve with reasonable effort), additional explanations are needed.

Our tack has been to study naturally occurring examples and to try to hypothesize rules that explain the pronoun use in those examples. To date we have concentrated our study on New York Times news articles. We hope to generalize some of our findings to other types of text genres as well.

Consider the following passage from the first several lines of one of the stories we analyzed:

Example 1:
When **Kenneth L. Curtis** was wheeled into court nine years ago, mute, dull-eyed and crippled, it seemed clear to nearly everyone involved that it would be pointless to put **him** on trial for the murder of **his** former girlfriend, Donna Kalson, and the wounding of her companion.

It had been a year since **Mr. Curtis** had slammed **his** pickup truck into them, breaking their legs. **He** then shot them both and, finally, fired a bullet into **his** own brain. **Mr. Curtis** lingered in a coma for months, then awoke to a world of paralysis, pain and mental confusion from which psychiatric experts said **he** would never emerge.

One expert calculated **his** I.Q. at 62.....

For convenience, we have indicated all references to the main character in bold.

A surprising thing to note about this passage is that not all of the anaphoric references to *Mr. Curtis* are pronouns even though he is arguably the focus of every sentence included. Previous work on pronoun generation would predict that a pronoun should be used if the same item remains in focus. Thus it appears that something other than a straightforward application of focusing or other pronoun resolution algorithms is necessary.

A second thing to note about this passage is that the sentences are generally long and complex and often contain several references to the same character. These types of sentences are very different from those that have been considered by any generation system that has rules for generating pronouns. In addition, it is not clear how most focusing or pronoun interpretation algorithms would handle them.

Looking for an explanation, one might turn to the underlying structure of the text. However, in doing so, care must be taken to ensure that the chosen structure both affects pronoun generation decisions and is something that is well-defined from a generation perspective. A structuring device such as paragraph breaks fails on both of these counts. At first one might expect to see a definite description at the beginning of a paragraph, and a pronoun inside a paragraph (as long as the item being referred to is in focus). While this correctly predicts the name at the beginning of paragraph two, notice there is a name in the middle of that same paragraph and a pronoun at the beginning of paragraph three. More importantly, from the perspective of a generation system, it is not at all clear how a paragraph break should be defined. Thus positing a

rule based on such a structuring device is not helpful. We must search for an answer that (1) explains the patterns of pronouns found in naturally occurring text, and (2) is based on information available to a sentence generation system.

In this paper we present our preliminary work in doing just that. Our work so far has led us to hypothesize several factors that affect pronoun generation and that are available to a sentence generation system. These factors include:

Sentence Boundaries – pronouns appear to be the preferred referring form for subsequent reference to an item within the same sentence.

Distance from Last Mention – when the last mention of an item is several sentences back in the text, a definite description is preferred.

Discourse Structure (as indicated by time change) – if we take time as a discourse structuring device, when the previous reference to an item is at the same time as the current reference, a pronoun is preferred. A definite description is preferred when the time has changed.

Ambiguity – potential ambiguity must be taken into account when choosing an anaphoric expression in that a pronoun should only be generated if it can be resolved correctly.

In the next section we discuss previous research on pronoun generation. This is followed by an introduction of time as a structuring device which affects pronoun generation. Next we investigate ambiguous anaphoric references. We follow this with an algorithm which decides when to use a pronoun versus a definite description when referring to some discourse entity. After this, we report on empirical results of the application of our algorithm to a corpus of New York Times articles.

Previous Work on Pronoun Generation

Few researchers have given serious consideration to the problem of pronoun generation. The most common factor considered has been the accessibility of the referent. For example, a pronoun would be generated if its referent was sufficiently prominent in the preceding text. Some early generation work (e.g., McDonald (1980), McKeown (1985), McKeown (1983), Appelt (1981)) used a simple rule to implement this idea based on focus (Sidner, 1979). The rule roughly stated that if the current sentence is about the same thing that the previous sentence was about, use a pronoun to refer to that thing. As was pointed out above, this rule does not provide a very good match with the referring expressions in our corpus.

In Dale (1992), Dale also discussed the generation of pronouns in the context of work on generating referring expressions (Appelt, 1985; Reiter, 1990). He specified an algorithm that essentially generated the smallest referring expression that distinguished the object in question from all others in the context.[1] He generated a pronoun (or ellipsis) if one were adequate and if the object being referred to was the center of the last utterance. As an example of the kinds of texts he generated consider: "Soak the butterbeans. Drain and rinse *them*."

[1]This algorithm was later revised in Dale & Reiter (1995) to more adequately reflect human-generated referring expressions and to be more computationally tractable.

Such an account of pronoun generation, based on center constancy appears to work quite well in the domain studied by Dale, but is unable to account for the patterns in other kinds of naturally occurring text.

The centering model (Grosz et al., 1995) itself makes predictions about pronoun generation only in the instances where Rule 1 is applicable. Recall that centering associates a set of forward looking centers with each utterance. This is an ordered list of the discourse entities introduced in the utterance. While an active area of research, in English the order is generally taken to be subject, object, other arguments to the verb in surface order, adjuncts of the verb in surface order. The backward looking center of an utterance is the highest ranked element of the previous utterance's forward looking centers that is realized in the current utterance. Centering's Rule 1 states that if any element of the previous utterance's forward looking centers list is realized in the current utterance as a pronoun, then the backward looking center must be realized as a pronoun as well (Grosz et al., 1995, p.214). Notice that the *Mr. Curtis* at the beginning of the second sentence in Example 1 is an apparent violation of this rule (because it is the subject of the previous sentence but is not pronominalized while another element mentioned in that sentence is). But, more generally, we must have a theory that is able to handle all cases of pronoun use.

A pronoun interpretation algorithm based on centering which relied on centering transition preferences was developed in Brennan et al. (1987). Using transition preferences in a pronoun generation rule would cover more cases of pronoun use than is covered by Rule 1, but the application of such transition preferences also proved unhelpful in explaining pronoun patterns in our corpus.

Grosz & Sidner (1986) and Reichman (1985) indicate that discourse segmentation has an effect on the linguistic realization of referring expressions. While this is intuitively appealing, it is unclear how to apply this to the generation problem (in part because it is unclear how to define discourse segments to a generation system). Passonneau (1996) argues for the use of the principles of information adequacy and economy in generating anaphoric expressions. Her algorithm takes discourse segmentation into account through the use of focus spaces which are associated with each discourse segment. Thus, Passonneau explains that a fuller description might be used at a discourse segment boundary because the set of accessible objects changes at such boundaries (though she combines this consideration with centering theory which may override the decisions due to segment boundaries). While Passonneau's algorithm seems quite appealing, notice that it provides no explanation of how a discourse segment (boundary) should be defined. The evaluation that she provided used the discourse segments provided by a set of naive subjects who indicated discourse segment boundaries in her texts. Without such boundaries provided, it is impossible to apply her algorithm. In some sense, the work presented here is consistent with Passonneau's theory. What we attempt to add is a definition of discourse segment which is well-defined and can be derived from input that any sentence generator must have in order to generate a sentence. On the other hand, we differ from Passonneau in that we do not attempt to make use of focus spaces in generation. Rather we

argue for evaluating informational adequacy on the basis of confusable objects near the current sentence in a discourse.

In the following section we hypothesize that discourse segment boundaries (or, perhaps, setting changes) do have an effect on appropriate anaphoric expression choice and that changes in time are markers of such boundaries (in the stories that we have analyzed). We hypothesize that if the current and previous reference to an entity occur in clauses referencing two different times, a definite description is used (and when the time referenced is the same, a pronoun is used).

Temporal Discourse Structure

In this section, we describe our approach to using discourse structure for choosing the right referring expression. Since we are working with stories from newspapers we were not able to identify the kind of discourse structure as assumed by Grosz & Sidner (1986), whose dialogues are more task-oriented and have clear intentional goals. We needed to find a structuring device that was both recognizable and part of the input to a sentence generation system. After investigating some work on narrative structure (Genette, 1980; Prince, 1982; Vogt, 1990), we concentrated on temporal structure. Temporal structure (an impoverished notion of the deictic center (Nakhimovsky, 1988; Wiebe, 1994) and related to temporal focus (Webber, 1988)) informally relates to the time being referenced in a text. This structure is often indicated by linguistic means and must be part of the input to a sentence generator. Changes in temporal structure or time may require world knowledge reasoning to recognize but are often indicated by either cue words and phrases (e.g., *"nine years ago", "a year", "for months", "several months ago"*), a change in grammatical time of the verb (e.g., past tense versus present tense), or changes in aspect (e.g., atomic versus extended events versus states as defined by Moens & Steedman (1988)).

In considering how time change might affect anaphoric expression choice, we consider the choice for the first mention of a discourse entity in a sentence where that entity has recently been referred to in the discourse. Our hypothesis is that: Changes in time reliably signal discourse segment boundaries in newspaper articles; definite descriptions should appear when the current reference to a discourse entity is in a different discourse segment from the last reference to that entity and pronouns should occur when the previous mention is in the same discourse segment.

Notice that our hypothesis does *not* cover *long distance* situations – where a discourse entity has not been referred to for several sentences; we have found that a definite description is almost always used in long distance situations regardless of time changes. A second place where time does not affect anaphoric expression choice is at very short distances – subsequent reference to an object *within the same sentence* is almost always realized as a pronoun regardless of whether or not the time has changed.

In order to evaluate this hypothesis, we mapped out the time being referenced in our texts on a clause-by-clause basis. For each clause in the texts, we indicated the time which was referred to. We distinguished between events that occurred at a single instance in time (atomic events) and events or states that occurred over a span of time (repeated atomic events, extended events, and states). For atomic events we allowed for both a specific time at which it occurred and for a non-specific time that indicated the range of uncertainty. We allowed time spans to have both specific end points and unspecified end points as well.

An example from our corpus with its associated temporal structure may illustrate these labels, the complexity of the texts under consideration, and how we propose pronoun generation is affected.

Example 2:

(47a) Questioned about the criminal activities of the football club,

(47b) **Mrs. Mandela** maintained

(47c) that **she** had never had any control over them.

(48a) This despite testimony from a half dozen former members

(48b) that they even had to get permission to go in and out of **her** yard.

(49a) **Mrs. Mandela** also said

(49b) **she** had disbanded the club

(49c) after **her** husband asked **her** to, despite evidence to the contrary.

(50) **Mrs. Mandela** faced questions from more than 10 lawyers representing various victims and the panel of commissioners and their investigators.

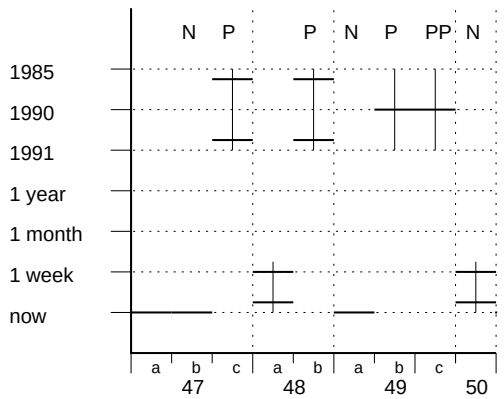

Figure 1: Temporal Structure for Example 2

Figure 1 contains the temporal structure for (each clause) of sentences 47-50 of one of our texts. The corresponding sentences (also broken into finite clauses are contained in Example 2.

Notice that sentence 47 consists of three clauses. The first two (47a and 47b) describe atomic events that are taking place at the "now" time of the story (during the proceedings against Mrs. Mandela). The third clause (47c) refers to an indefinite span of time in the past (roughly from 1985 to 1991, during which Mrs. Mandela's football club existed). Note, the use of the past perfect indicating the change in time and setting.

As Figure 1 illustrates, there is a name (N) reference to Mrs. Mandela in (47b), and a pronominal (P) reference to her in (47c). This pronoun is used even though there is a change in time between (47b) and (47c) and is explained because this is the second reference to Mrs. Mandela within the same sentence (a condition that overrules the time change hypothesis).

(48a) represents a change back to the time of the proceedings (note the discourse-deictic reference (Webber, 1991) "This"). In this case, a time span indicates multiple atomic events occurring during that time. (48b) again points to the time span in the past, though this is not explicitly marked linguistically as it was in (47c). Here world knowledge must be used to understand the time referenced in this clause. Note, however, that the time would have to be part of the input to a generator, and thus our rules are completely well-defined from the generation perspective.

The use of a pronoun to refer to Mrs. Mandela in (48b) is warranted by our hypothesis, because the clause containing the previous mention of Mrs. Mandela, (47c), references the same time as is referenced in (48b).

Because there is a time change between (48b) and (49a), our hypothesis explains the appearance of the proper name in (49a) even though it occurs just after a pronoun (in (48b)) co-specifying with the same character.[2] The pronouns in the remainder of (49) are explained because they are subsequent references within the same sentence (despite the fact that they refer to an unspecified time in the past which is different from the time referenced in (49a)). Finally, the use of a name in (50) is again indicated by the change in time between (49c) and (50).

Ambiguities

Of course, the choice of referring expression is not only guided by discourse structure; there is also an influence due to ambiguities. Dale (1992) generated referring expressions so that they could be distinguished from the other discourse entities mentioned in the context. This strategy can be interpreted as: Generate a pronoun whenever it is not ambiguous. However, how one should define context is not quite clear. For this definition we choose a span of text considered important in our previous work on anaphora resolution (Strube, 1998), and define a referring expression as ambiguous if there is a competing antecedent (i.e. another discourse entity matching in number and gender) mentioned in the previous sentence or to the left of that referring expression in the current sentence. Of the 437 referring expression in the texts we analyzed, 104 were considered ambiguous by this definition. Of these only 51 were realized as a definite description. Thus a rule which says to use a definite description if a pronoun would be ambiguous according to the definition appears to be too strict. Therefore we need to consider ambiguous cases in more detail.

To handle these cases we turned to pronoun resolution algorithms. Our intuition was that a pronoun could be generated to refer to a particular discourse entity if a pronoun resolution algorithm would choose that entity as a referent for the pronoun. To our knowledge, there are only two pronoun resolution algorithms that are specified in enough detail to work on unrestricted naturally occurring text: Brennan et al. (1987) using the definition of utterance according to Kameyama (1998), and Strube (1998). Strube (1998) evaluated the effectiveness of these two algorithms on the task of pronoun resolution in some naturally occurring texts. Because Strube's algorithm showed significantly better results, we have turned to it for guidance in pronoun generation.

The idea is that if we want to refer to a discourse entity, `E`, but there is a competing antecedent, `C`, we look to Strube's algorithm in the following way. If Strube's algorithm would resolve a pronoun to be `E`, we use a pronoun. If, instead, Strube's algorithm would prefer `C` as the referent of the pronoun, we use a definite description to refer to `E`.

Our analysis showed that the use of Strube's algorithm showed improvement, but it seemed to be too liberal with suggesting pronouns when the competing antecedent was in the previous sentence. Thus, the rule that we settled on is shown in Figure 2.

1. *If* this is the first occurrence of X in the current sentence and
   (a) *if* there is a competing antecedent in the *previous* sentence, use a definite description;
   (b) *if* there is a competing antecedent in the *same* sentence (i.e., to the left) and
      i. *if* Strube's algorithm would resolve a pronoun in this position to be X, use a pronoun;
      ii. *else* use a definite description;
2. *if* this is a subsequent occurrence of X in the current sentence and
   (a) *if* there is an intervening competing antecedent, use a definite description;
   (b) *if* there is no intervening competing antecedent, use a pronoun.

Figure 2: Realization of the Referring Expression X when Competing Antecedents Exist

Anaphoric Expression Generation Algorithm

In the previous sections we have argued that temporal structure should influence pronoun choice and that ambiguous cases need to be handled separately. In addition, we found that if the distance between the current reference and the previous mention of the discourse entity being referred to was far (over two sentences), a definite description was almost always used. Finally, we found that when an entity was referred to multiple times in a sentence, a pronoun was usually used for subsequent references in a sentence.

Based on these findings, we propose the algorithm for realizing anaphoric expressions shown in Figure 3.

Empirical Data

We applied the algorithm described in the previous section to three texts from the *New York Times*. Articles were ranging from a frontpage article to local news. We applied the algorithm to all references to persons in these texts. The algorithm was correct in 370 cases (84.7%), and wrong in 67 cases (15.3%). In Figure 4 we show the distribution over the rules specified in the algorithm.

In order to interpret the results of the algorithm, we must have some comparison. We use the simple scheme shown in Figure 5 for comparison purposes.

[2] Note that the use of this definite description cannot be explained by a topic shift since there is no topic shift in between the previous text and (49). At least two discourse entities (*"Mrs. Mandela"* and the *"football club"*) are constant, only Mrs. Mandela's husband does not occur in the immediately preceding sentences.

1. *If* this is a **long distance anaphoric reference** (i.e., if the previous reference to X was more than two sentences prior) use a **definite description**;

2. *else*
   if this is an **unambiguous reference** (i.e., there is no competing antecedent) and this is an **intra-sentential anaphor** (i.e., this is not the first mention of X in the current sentence) use a **pronoun**;

3. *else*
   if this is a **time change** (i.e., there is a difference between the time of the clause with the previous reference to X and the time of the current clause) use a **definite description**;

4. *else*
   if there is a **competing antecedent** (i.e., another object in the previous or current sentence that matches the type and number of X) use the **rule found in Figure 2**;

5. *else*
   for the **remaining cases** use a **pronoun**.

Figure 3: Algorithm for Generating the Appropriate Form for a Referring Expression X

Rule Name		number	percentage
All Rules	correct	370	84.7%
	wrong	67	15.3%
Long Distance Anaph.	correct	46	97.9%
	wrong	1	2.1%
Intra-sentential Anaph.	correct	168	96%
	wrong	7	4%
Time Rules (3 & 5)	correct	116	72.5%
	wrong	44	27.5%
Competing Antecedent	correct	40	72.7%
	wrong	15	27.2%

Figure 4: Results of the Algorithm

The results of applying these rules give 343 correct cases (78.5%) and 94 incorrect ones (21.5%). Hence our algorithm reduces the error rate by 28.9%.

Related Research

A significant amount of work in linguistics has investigated the use of different kinds of anaphoric referring expressions in discourse and their relationship to ease of comprehension. See Arnold (1998) for a discussion of several of the factors involved in referring expression choice. In many cases the various factors seem to affect the *accessibility* of a referent (where accessibility is intended in a broad sense to cover both "topic accessibility" (Givon, 1983) and accessibility due to factors such as recency of mention). Basically, the more accessible a referent the more underspecified a referring expression should be. Accessibility explains the apparent "name-name penalty" as examined in Gordon & Hendrick (1998), for example.

Our work argues that factors beyond accessibility must be considered in anaphoric expression choice. It is consistent with work such as Vonk et al. (1992) whose experiments indicate that a referring expression "... that is more specific than is necessary for the recovery of the intended referent ... marks the beginning of a new theme concerning the same dis-

1. *If* this is a **long distance anaphoric reference** (i.e., if the previous reference to this item is greater than two sentences prior) use a **definite description**;

2. *else*
   if there is a **competing antecedent** (i.e., another object in the previous or current sentences that matches the pronoun which would be used to refer to this entity) use a **definite description**;

3. *else*
   the anaphoric expression would be **unambiguous** so use a **pronoun**.

Figure 5: Simple Algorithm

course referent." (Vonk et al., 1992, page 304). They argue that such overspecified expressions are serving a discourse function of indicating boundaries. This work does not define what a discourse segment boundary actually is. On the other hand, using the definition of time change as a boundary condition, our work is consistent with their hypothesis. Interestingly, Vonk et al. (1992) found that in discourses where a theme change was well marked by other means (e.g., by a preposed adverbial phrase or a subordinate clause indicating time or place) that pronouns were much more common even though a new theme was begun. Presumably such phrases mark the theme change well, and thus it is not necessary to also mark the change via an overspecified description. We are currently reanalyzing our data in light of that finding to see of it provides a fuller account of the naturally occurring data.

Approaches which define discourse segments on the basis of reference resolution (Sidner, 1979; Suri & McCoy, 1994; Strube & Hahn, 1997) are not useful for our purposes because they require referring expressions for recognizing segment boundaries. In contrast to these approaches, we define segment boundaries independently from reference resolution so that in this respect our work is in line with Grosz & Sidner's (1986) definitions.

Future Directions

In analyzing our data, there are several places for further consideration. One problem is that our rule which indicates a definite description should be used in a time change overgenerates definite descriptions. Following Vonk et al. (1992) we plan to investigate whether definite descriptions might best be viewed as boundary markers and whether other markers of discourse boundaries (e.g., preposed adverbial phrases) are found in places where our algorithm suggests a definite description because of a time change but a pronoun appears in the text.

In places where our algorithm overgenerates pronouns, a more sophisticated time analysis may be helpful. Our current analysis distinguishes between four types of time and is driven by both semantic cues in the text (e.g., adverbial time phrases) and changes in tense. Nakhimovsky (1988) also uses changes in "time scale" as a marker for changes in time. We plan to investigate this to see whether it explains more of the examples. Nakhimovsky (1988) also describes several other markers for a setting change, and these will also be investigated to see if they are indicative of definite description use.

Another line of future research involves further investiga-

tion of the ambiguous cases. Our current rule was developed by evaluating several different possibilities (e.g., using time change rules, different pronoun resolution algorithms) and selecting a rule that explains most of the cases. Still, the number of ambiguous cases is fairly small and analyzing more texts and concentrating on cases where the current rule makes an incorrect prediction may lead us to a more robust rule.

Conclusions

Pronouns occur frequently in texts and have been hypothesized to play a large role in text coherence. Yet, pronoun generation has not been studied in detail. If future natural language generation systems are to produce coherent, natural texts, they must use rules for generating pronouns that produce pronouns in roughly the same places that human-produced texts do. Moreover, the rules must be based on information that would be available to a sentence generator. At the same time, in order to evaluate rules, they must be based on information that can be gleaned from a text.

In this work we have argued that changes in setting, as indicated by changes in time, provide an explanation for patterns of pronoun use in naturally occurring text. That is, even in places where a pronoun would be unambiguous, a definite description might be used when the time of the sentence is different from the time of the sentence in which the previous mention was made. This hypothesis provides an explanation for many of the uses of definite descriptions found in the studied texts. Other uses of definite descriptions occur because of ambiguities. We have suggested a rule which addresses when such ambiguities should not preclude the generation of a pronoun. Our scheme appears to be a reasonable explanation for the patterns of pronoun use found in our corpus.

Acknowledgments. This work was done while the first author was visiting the Institute for Research in Cognitive Science and while the second author was a post-doctoral fellow there (NSF SBR 8920230). We would like to thank Jennifer Arnold and the centering group at UPenn for helpful discussions. We would also like to thank the anonymous reviewers.

References

Appelt, D. E. (1981). *Planning Natural-Language Utterances to Satisfy Multiple Goals*, (Ph.D. thesis). Stanford University. Also appeared as: SRI International Technical Note 259, March 1982.

Appelt, D. E. (1985). Planning English referring expressions. *Artificial Intelligence*, 26(1):1–33.

Arnold, J. E. (1998). *Reference Form and Discourse Patterns*, (Ph.D. thesis). Stanford University, Department of Linguistics.

Brennan, S. E., M. W. Friedman & C. J. Pollard (1987). A centering approach to pronouns. In *Proceedings of the 25th Annual Meeting of the Association for Computational Linguistics,* Stanford, Cal., 6–9 July 1987, pp. 155–162.

Dale, R. (1992). *Generating Referring Expressions: Constructing Descriptions in a Domain of Objects and Processes*. Cambridge, Mass.: MIT Press.

Dale, R. & E. Reiter (1995). Computational interpretations of the Gricean maxims in the generation of referring expressions. *Cognitive Science*, 18:233–263.

Genette, G. (1980). *Narrative Discourse: An Essay in Method*. Ithaca, N.Y.: Cornell University Press.

Givon, T. (1983). Topic continuity in spoken English. In T. Givon (Ed.), *Topic Continuity in Discourse: A Quantitative Cross-Language Study*. Amsterdam, Philadelphia: John Benjamins.

Gordon, P. C. & R. Hendrick (1998). The representation and processing of coreference in discourse. *Cognitive Science*, 22(4):389–424.

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

Grosz, B. J. & C. L. Sidner (1986). Attention, intentions, and the structure of discourse. *Computational Linguistics*, 12(3):175–204.

Kameyama, M. (1998). Intrasentential centering: A case study. In M. Walker, A. Joshi & E. Prince (Eds.), *Centering Theory in Discourse*, pp. 89–112. Oxford, U.K.: Oxford University Press.

McDonald, D. D. (1980). *Natural Language Production as a Process of Decision Making Under Constraint*, (Ph.D. thesis). MIT.

McKeown, K. R. (1983). Focus constraints on language generation. In *Proceedings of the 8th International Joint Conference on Artificial Intelligence,* Karlsruhe, Germany, August 1983, pp. 582–587.

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

Moens, M. & M. Steedman (1988). Temporal ontology and temporal reference. *Computational Linguistics*, 14(2):15–28.

Nakhimovsky, A. (1988). Aspect, aspectual class, and the temporal structure of narrative. *Computational Linguistics*, 14(2):29–43.

Passonneau, R. (1996). Using centering to relax Gricean constraints on disocurse anaphoric noun ophrases. *Language and Speech*, 39(2):229–264.

Prince, G. (1982). *Narratology: The Form and Functioning of Narrative*. Berlin: Mouton.

Reichman, R. (1985). *Getting Computers to Talk like You and Me*. Cambridge, Mass.: MIT Press.

Reiter, E. (1990). Generating descriptions that exploit a user's domain knowledge. In R. Dale, C. Mellish & M. Zock (Eds.), *Current Research in Natural Language Generation*. London: Academic Press.

Sidner, C. L. (1979). *Towards a Computational Theory of Definite Anaphora Comprehension in English*. Technical Report AI-Memo 537, Cambridge, Mass.: Massachusetts Institute of Technology, AI Lab.

Strube, M. (1998). Never look back: An alternative to centering. In *Proceedings of the 17th International Conference on Computational Linguistics and 36th Annual Meeting of the Association for Computational Linguistics,* Montréal, Québec, Canada, 10–14 August 1998, Vol. 2, pp. 1251–1257.

Strube, M. & U. Hahn (1997). Centered segmentation: Scaling up the centering model to global referential discourse structure. In *Proceedings of the 19th Annual Conference of the Cognitive Science Society,* Palo Alto, Cal., 7–10 August 1997.

Suri, L. Z. & K. F. McCoy (1994). RAFT/RAPR and centering: A comparison and discussion of problems related to processing complex sentences. *Computational Linguistics*, 20(2):301–317.

Vogt, J. (1990). *Aspekte erzählender Prosa: Eine Einführung in Erzähltechnik und Romantheorie* (7th ed.). Opladen: Westdeutscher Verlag.

Vonk, W., L. G. Hustinx & W. H. Simons (1992). The use of referential expressions in structuring discourse. *Language and Cognitive Processes*, 7(3/4):301–333.

Webber, B. L. (1988). Tense as discourse anaphor. *Computational Linguistics*, 14(2):61–73.

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

Wiebe, J. M. (1994). Tracking point of view in narrative. *Computational Linguistics*, 20(2):233–287.