Open Research Online

The Open University's repository of research publications
and other research outputs

The Influence of layout on the interpretation of referring expressions

Book Section

For guidance on citations see FAQs.

oro.open.ac.uk

The Influence of Layout on the Interpretation of Referring Expressions[1]

Nadjet Bouayad-Agha, Donia Scott and Richard Power
ITRI, University of Brighton
Firstname.Surname@itri.brighton.ac.uk

1 Introduction

The division of text into visual segments such as sentences, paragraphs and sections achieves many functions, such as easing navigation, achieving pragmatic effect, improving readability and reflecting the organisation of information (Wright, 1983; Schriver 1997). In this paper, we report a small experiment that investigates the effect of different layout configurations on the interpretation of the antecedent of anaphoric referring expressions.

Layout has so far played little role in Natural Language Generation (NLG) systems. The layout of output texts is generally very simple. At worst, it consists of only a single paragraph consisting of a few sentences; at best it is predetermined by schemas (Coch, 1996; Porter and Lester, 1997) or discourse plans (Milosavljevic, 1999). However, recent work by Power (2000) and Bouayad et al. (2000) has integrated graphically signalled segments (e.g., by whitespace, punctuation, font and face alternation) such as paragraphs, lists, text-sentences and text-clauses in a hierarchical tree-like representation called the *document structure*.[2] This work was carried out within the ICONOCLAST project (Integrating CONstraints On Layout and Style), which aims at automatically generating formatted texts in which the formatting decisions affect the wording and vice-versa.[3]

If document structure affects the comprehensibility of referring expressions, this must be taken into account in any attempt to generate felicitous formatted texts. This will go a step further from current research in the automatic generation of referring expressions, where only the effect of discourse structure and grammatical function has been investigated (Dale and Reiter, 1995; Cristea et al., 1998; Walker et al., 1998; Kibble and Power, 1999).

Document designers and professional writers follow the general principle that the structure of the message should be reflected onto the page (Wright, 1983; Schriver, 1997). College students are also advised to follow this rule when writing expository texts (Ostorm and Cook, 1988). In the research literature, text linguists often use the notion of semantic (or topical) paragraph (Hinds, 1977; VanDijk, 1977; Longacre, 1979; Brown and Yule, 1983), assuming at best that orthographic and semantic paragraph coincide, and at worst that the orthographic paragraph has just a purely esthetic role (Longacre, 1979}.[4] Researchers in text comprehension have also experimentally demonstrated the role of paragraphing in understanding the structure of the message (Kieras, 1978; Bond and Hayes, 1984; Stark, 1988; Goldman, 1995). Two main approaches are used in these experiments: either de-paragraphing a text and asking subjects to re-paragraph it, or manipulating

---

[2] Following Nunberg (1990), a text-sentence in expository texts is an orthographic sentence starting with a capital and ending with a full stop. A text-sentence can consist of two text-clauses separated by a semi-colon.

[4] The paragraph has been the object of most research efforts regarding the organisation of information, probably because it corresponds to a well-recognised unit of meaning.

the paragraph-initial sentence so that it does not always coincide with the main idea of that paragraph. Although the results show the impact of layout on interpretation, they are not the product of an explicit manipulation of an abstract layout representation such as the document structure and therefore do not illustrate the effect of different layout decisions on interpretation.

The decisions on how to visually structure the information in a text has an impact on the cohesive devices of language (Halliday, 1976) such as referring expressions. This comes as no surprise if we consider the reasons for dividing a text using a paragraph break: it "could serve to alert the reader to look for argument overlap in the global-passage context rather than merely in the local […] context" (Goldman et al, 1995, p.277), or to signal to readers to free their working memory (Hofmann, 1989). This means that the degree of accessibility of the antecedent of an anaphoric expression is affected by paragraphing, alongside other factors such as distance and saliency (Ariel, 1990). Ariel (1990) provides some evidence for this claim by analysing in a small corpus of newspaper texts the difference in distribution of the three main forms of anaphoric referring expressions according to their position in the text. She observes the following continuum: over 80% of pronouns occur in the same sentence or a contiguous sentence with their antecedent, around 80% of demonstratives occur in a contiguous or a non-contiguous sentence within the same paragraph, and over 80% of definite descriptions occur in non-contiguous sentence within the same paragraph or in different sentences across paragraphs. Although the effect of distance is not neutralised, these differences in distribution provide some support to the effect of location on the form of the anaphor. The different locations reflect in turn a different degree of topical continuity. Bond and Hayes (1984) also indirectly prove the impact of layout on referring expressions by showing that subjects use sentence-initial noun phrases, but not pronouns, as a cue for paragraphing a text stripped of its paragraph breaks (Bond and Hayes, 1984).

In this experiment, we manipulate the document structure representation of texts that contain an anaphoric expression and two potential antecedents. These manipulations affect the accessibility of these potential antecedents, thereby determining the choice of one over the other.

2 Design and Materials

The materials used in this experiment are short texts consisting of four consecutive clauses (C1 to C4). These four clauses have the following function:

- C1 introduces the discourse, thus subsuming the rest of the argument. This is done so that when using a list, C1 dominates it.

- C4 is the target clause, limited from 7 to 9 words. It contains the anaphor (ANA). ANA is a demonstrative noun phrase of the form "these + NOUN". Demonstrative noun phrases (NPs) were chosen because they are flexible regarding their location relative to their antecedent, as noted by Ariel (1990). She observed that demonstrative NPs tend to be positioned with their antecedent in roughly the same proportion in contiguous sentences, non-contiguous sentences and contiguous paragraphs, thus referring to them as intermediate accessibility markers. Thus, demonstrative NPs seem like the ideal anaphoric form for studying the influence of layout on the interpretation of anaphors, although they are the least studied anaphoric form.

- C2 and C3 each contain a noun phrase referred to as NP_{C2} and NP_{C3} respectively. The antecedent of ANA can either correspond to the (sum of) entities referred to by NP_{C3} or to the sum of entities referred to by NP_{C2} and NP_{C3}.

The materials were varied along two dimensions of the document structure:

- **Categorical,** that is, the text-level categories of the nodes in the document structure. We call this the *layout* condition. It has three possible values: text-sentence, paragraph and list item.

- **Structural,** that is, whether C3 and C4 are grouped under the same text-level node or under a different text-level node. We call this the *sameness* condition. It has two possible values: same or different.

An example for each of the six possible versions (3 *layout* conditions * 2 *sameness* conditions) of the same text is given in figure 1 below. For the sake of clarity, we highlight in the examples the anaphor and its possible antecedents; this of course was not signalled to the participants. A simplified document structure representation for each of the six versions is given in figure 2. As the figure illustrates, the "Different" condition produces a balanced document structure, in which C2, C3 and C4 are siblings, whereas the "Same" condition produces an unbalanced document structure, in which C2 is not the sibling of C3 and C4.

In what follows, a detailed specification of the six possible configurations is given:

- **Sentence:** In the "Same" condition, C3 and C4 are in the same orthographic sentence, separated by a semi-colon. In the "Different" condition, C3 and C4 are two consecutive orthographic sentences, either displayed on the same line or on justified lines so as to avoid the impression that there was a deliberate line break after C3.

- **Paragraph:** C1 and C2 are presented in different paragraphs separated by a blank line. All but the first paragraph are indented. In the "Same" condition, C3 and C4 are presented as two text-sentences within the same paragraph. In the "Different" condition, C3 and C4 are presented in different paragraphs separated by a blank line.

- **List:** C1 is presented as single paragraph, followed by a blank line, followed by the list whose items are introduced by hyphen style bullet points and separated by blank lines.

Our first prediction is that the preferred antecedent of ANA in the "Same" condition corresponds to the entities referred to by NP_{C3} whereas the preferred antecedent of ANA in the "Different" condition corresponds to the sum of entities from NP_{C2} and NP_{C3}. We refer to the first interpretation as "Last" and to the second interpretation as "Both".

Thus, in the examples presented above, the preferred antecedent of "these effects" in the "Same" configurations should be NP_{C3}, that is, *a small amount of hair discolouration and a skin rash*. In the ``Different'' configurations, the antecedent should the sum of entities in NP_{C2} and NP_{C3}, that is, *headache and stomach upset* and *a small amount of hair discolouration and a skin rash*.

Same, Sentence

Some side effects may occur while you are taking your medicine. Common effects include *headache and stomach upset*. You may also get *a small amount of hair discolouration and a skin rash*; you can overcome <u>these effects</u> by lowering your dose.

Different, Sentence

Some side effects may occur while you are taking your medicine. Common effects include *headache and stomach upset*. You may also get *a small amount of hair discolouration and a skin rash*. You can overcome <u>these effects</u> by lowering your dose.

Same, Paragraph

Some side effects may occur while you are taking your medicine.

Common effects include *headache and stomach upset*.

You may also get *a small amount of hair discolouration and a skin rash*. You can overcome <u>these effects</u> by lowering your dose.

Different, Paragraph

Some side effects may occur while you are taking your medicine.

Common effects include *headache and stomach upset*.

You may also get *a small amount of hair discolouration and a skin rash*.

You can overcome <u>these effects</u> by lowering your dose.

Same, List

Some side effects may occur while you are taking your medicine.

− Common effects include *headache and stomach upset*.

− You may also get *a small amount of hair discolouration and a skin rash*. You can overcome <u>these effects</u> by lowering your dose.
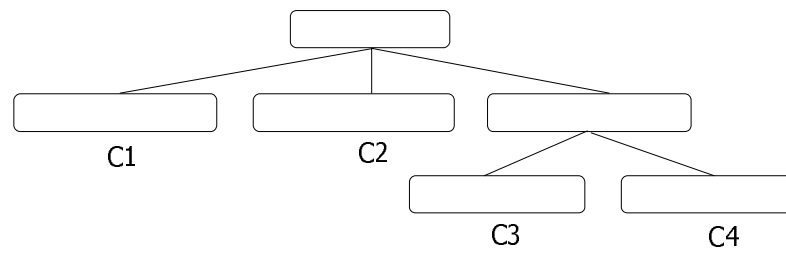
Different, List

Some side effects may occur while you are taking your medicine.

− Common effects include *headache and stomach upset*.

− You may also get *a small amount of hair discolouration and a skin rash*.

− You can overcome <u>these effects</u> by lowering your dose.
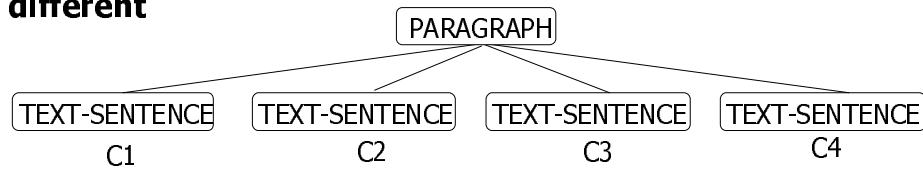
Figure 1: Example of a text in the six possible versions
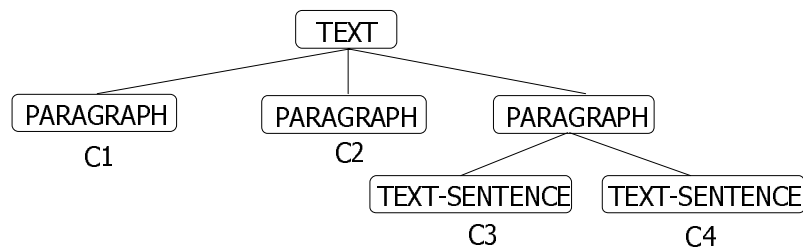*(same/different * text-sentence/paragraph/list)*
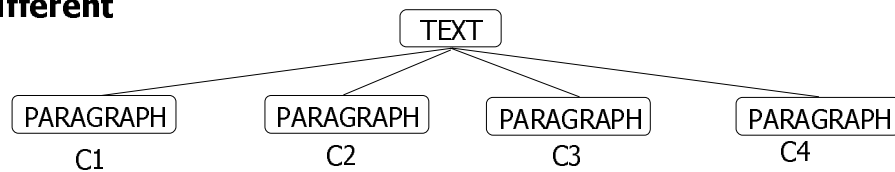
sentence
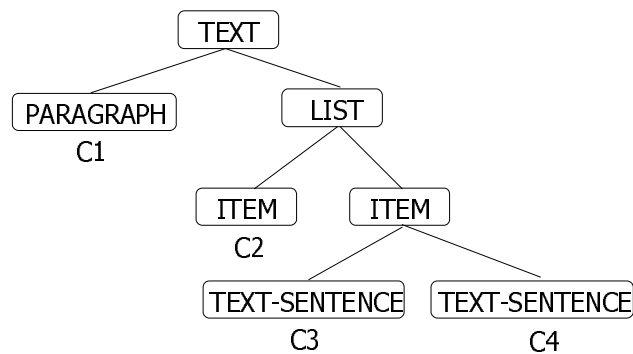
same

different

paragraph
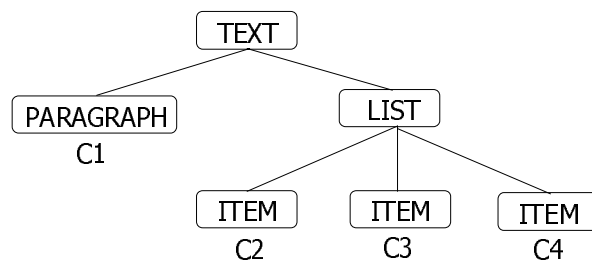
same

different

list

same

different

Figure 2: Simplified document structure for each of the six text versions

Our second prediction is that the processing of the antecedent of the anaphor should take longer in the "Different" configuration than in the "Same" configuration. The rationale behind that is that, although the distance between the relevant noun phrases stays the same, a layout division may prompt the reader to integrate the current information with what has gone before, as hypothesised by Hofmann (1989) and Goldman et al (1995).

In order to reinforce the legitimacy of using a demonstrative NP as the anaphoric form in all layout contexts, we used the following constraints when constructing the texts. Firstly, the head noun of the anaphor must be different from the head nouns of both NP_{C2} and NP_{C3}, thus avoiding the effect of word repetition. Note however that this may not prevent conceptual rather than lexical identity to reactivate a potential antecedent in memory even when the antecedent is distant from its anaphor (O'Brien, 1997).

The lexical requirement on the demonstrative anaphor entails that it assumes a role of redescription (Gundel, 1993). We believe that this increases the legitimacy of using a demonstrative form, particularly in short distance reference such as within the same sentence. However, it also means that NP_{C3} needs to be complex enough in order to legitimise the use of a demonstrative of redescription when the antecedent is in C3. Therefore, as well as being a plural noun phrase, NP_{C3} is always a coordination of singular nouns or modifiers (e.g., *a small amount of hair discoloration and a ski rash*). The last noun in the coordination is always singular to avoid the possibility that the plural ANA might refer back to that last noun only. The conjunction of coordination used is "and" as it is the most likely form for plural reference (Sanford, 1990).

NP_{C2} and NP_{C3} are always in the same grammatical role in order to facilitate the summing of their references. Finally, NP_{C2} and NP_{C3} are in non-subject position, either in direct or indirect object position. Indeed, Gundel et al (1993) claim that the antecedent of a demonstrative need not be in focus; it only needs to be in the list of preferred centers, that is, it needs to be activated. If it were in focus, then a pronoun might be preferred. Similarly, ANA is in a non-subject position in order to avoid the interpretation of topic shift from one clause to the next (i.e., from C3 to C4), thus favouring NP_{C3} as the antecedent.

3 Method

Twenty-four expository texts of various genres (e.g., instructions, reports) were constructed according to the specifications and design presented above. One hundred and eight text versions, that is, 18 texts each in six different versions, were distributed in six different sets. The six remaining texts, each in a different version, were used in a practice session prior to the experiment.

Twenty-four psychology students from the University of Sussex took part in the 30-minute experiment for a £5 payment. They were naive as to the purpose of the experiment. Each participant was presented with all 18 text versions in a random order. At the end of each text, a multiple-choice question was presented to the participant (e.g., *Which effects can be overcome by lowering your dose?*), the answer to which is either the entities in NP_{C3} or the sum of entities in NP_{C2} and NP_{C3}.

Reading time was monitored with a tachistoscope. Each clause was presented individually and maintained on the screen until the participant pressed a button, upon which the next clause appeared, thus gradually revealing the structure of the entire text. The reading time for C4 was calculated as the time between pressing the button for C4 to appear and pressing the button again to

signal that C4 was understood. At this point, the participant could turn to the answer sheet. The gradual presentation of the structured text, although unnatural especially when formatted text is involved, ensures that the interpretation is incremental.

4 Results and Discussion

Table 1 illustrates the results of the interpretation experiment in the "Same" and "Different" configurations regardless of the layout conditions. As expected, the results are highly significant, with $\chi^2=52.943$ (p<0.0001). The preferred antecedent varies depending on the layout conditions. In particular, there is a strong preference for interpreting the sum of entities in both NP_{C2} and NP_{C3} as the antecedent of the anaphor in the "Different" condition (82% vs 18%). On the other hand, the antecedent is chosen almost equally as the last or both NPs in the "Same" condition (51.4% vs 48.6%). Thus, overall, the texts strongly favour the sum of entities in both clauses as the preferred interpretation regardless of layout, as illustrated by the total percentage distribution (65.3% vs 34.7%). This shows that readers use layout as a cue to help interpretation, but that they also strongly rely on semantic relations. Similar findings were reported by Stark (1988).

Nevertheless, the different layout configurations still have an effect on interpretation if, instead of looking at the overall distribution, we look at the distribution of each interpretation across the two layout conditions. We find that the "Last" interpretation is given in 74% of cases in the "Same" configuration versus only 26% in the "Different" configuration. Similarly, the "Both" interpretation is given in 62.7% of cases in the "Different" configuration versus 37.3% in the "Same" configuration.

	Same		Different		Total	
Last	111	51.4%	39	18%	150	34.7%
Both	105	48.6%	177	82%	282	65.3%
Total	216	100%	216	100%	432	100%

Table 1: Results of the interpretation experiment for all layout conditions

Table 2 shows the results for each layout condition. The results are highly significant for all three layout conditions (p<0.0001): paragraph $\chi^2=20.03$, list $\chi^2=17.099$, sentence $\chi^2=16.061$. The percentage distribution in the "Same" and "Different" conditions in the three layout conditions follow the same tendency as for the global results, with the exception of the "Same" condition in the list configuration, where the "Both" interpretation is slightly higher than the "Last" interpretation.

	Sentence			Paragraph			List		
	Same	Different	Total	Same	Different	Total	Same	Different	Total
Last	37 (51.4%)	14 (19.4%)	51 (35.4%)	40 (55.6%)	14 (19.4%)	54 (37.5%)	34 (47%)	11 (15.3%)	45 (31.25%)
Both	35 (48.6%)	58 (80.6%)	93 (64.6%)	32 (44.4%)	58 (80.6%)	90 (62.5%)	38 (53%)	61 (84.7%)	99 (68.75%)

Table 2: Results of the interpretation experiment for each layout condition

The calculation of the Fisher Exact Probability Test (FEBT) for each text shows that 10 out of 18 texts match significantly with our predictions.[5] The calculation of the FEBT for each participant shows that 7 out of 24 participants follow our predictions, 4 of which with $p<0.001$. Within the 17 remaining participants, 4 had the policy of choosing the sum of NP_{C2} and NP_{C3} as the antecedent, regardless of the presentation (same/different).

There was no significant difference between the reading times in the "Same" and "Different" conditions. One explanation to this might be that the layout division in the "Different" condition does not imply a change in topic and therefore, the reader is not required to try relating what he is reading with what he has read much earlier in the text. Similarly, Clark and Sengul (1979) found no significant difference in reading ttime when the antecedent of an anaphor was located in the previous orthographic sentence or in the same orthographic sentence as its anaphor.

5 Conclusion

The results of the interpretation experiment provide strong support for our claim. The overall results show that, when the document structure is balanced, subjects have a strong preference for taking the sum of entities in the two text segments as the plausible antecedent of the anaphor. On the other hand, there is no preferred interpretation when the document structure is unbalanced, with subjects choosing almost equally between the entities in the two text segments or in the last text segment as the preferred antecedent. We may advance the hypothesis that the default interpretation regardless of layout is the sum of entities in the two clauses preceding the anaphor. When layout is introduced, a balanced document structure reinforces the default interpretation whereas an unbalanced document structure may hinder it or even favour another interpretation, as suggested by the results of the experiment per text or per subject.

These findings show that layout has to be considered as an important factor for the generation of unambiguous referring expressions. Indeed, a document structure in which the groupings reflect the semantic relations minimise the risks of ambiguity. This issue may be particularly important in corpora that use layout and back reference extensively, such as the Patient Information Leaflet corpus we used as the basis of the ICONOCLAST project, and whose observation triggered this research.[6]

Clearly, further research is needed in order to understand better the interaction between layout and referring expressions and other generation decisions. For example, there may be situations where other constraints prevent the document structure from reflecting the semantic groupings. In this case, different linguistic devices may be used for preventing misinterpretation such as grammatical roles or anaphoric form.

---

[5] Unless otherwise specified, the level of significance is taken to be the standard one ($p<0.05$).
[6] A typical Patient Information presents series of ingredients, warnings or side-effects, grouped in paragraphs and/or vertical lists, with back reference to some or all of these entities.

References

Argyrous, G. 1997. *Statistics for Social Science*, MacMillan Press Ltd.

Ariel, M. 1990. *Accessing Noun-Phrase Antecedents*, Routledge.

Bond, S.J. and Hayes, J.R. 1984. Cues people use to paragraph text. *Research in the Teaching of English*, 18(2):147-67.

Bouayad-Agha, N., Power, R. and Scott, S. 2000. Can text structure be incompatible with rhetorical structure? In *Proceedings of the International Natural Language Generation Conference*, pages 194-200, Mitzpe Ramon, Israël, 12-16 June.

Brown, G. and Yule, G. 1983. *Discourse Analysis*. Cambridge Textbooks in Linguistics.

Campbell, K.S. 1995. *Coherence, continuity and cohesion - theoretical foundations for document design*, Lawrence Erlbaum Associates.

Clark, H.H. and Sengul, C.J. 1979. In search of referents for nouns and pronouns. *Memory and Cognition*, 7(1):35-41.

Coch, J. 1996. Overview of Alethgen. In *Demonstrations and Posters of the Eighth International Natural Language Generation Workshop (INLG'96)*, Herstmonceux Castle, Sussex, UK.

Cristea, D., Ide, N. and Romary L. 1998. Veins Theory: An approach to global cohesion and coherence. In *Proceedings of Coling/ACL*, Montréal.

Dale, R. and Reiter, E. 1995. Computational interpretations of the gricean maxims in the generation of referring expressions. *Cognitive Science*, 19:233-263.

Goldman, S.R., Saul, E.U. and Coté, N. Paragraphing, Reader, and Task Effects on Discourse Comprehension. *Discourse Processes*, 20:273-305.

Gundel, J.K., Hedberg, N. and Zacharski, R. 1993. Cognitive status and the form of referring expressions in discourse. *Language*, 69(2).

Halliday, M.A.K. and Hasan, R. 1976. *Cohesion in English*. Longman.

Hinds, J. 1977. Paragraph Structure and Pronominalization. *Papers in Linguistics*, 10:77-97.

Hofmann, T.R. 1989. Paragraphs & anaphora. *Journal of Pragmatics*, 13:239-250. North Holland.

Kibble, R. and Power, R. 1999. Using centering theory to plan coherent texts. In *Proceedings of the 12th Amsterdam Colloquium*, pp. 187-192, Amsterdam. The Netherlands.

Kieras, D.E. 1978. Good and Bad Structure in Simple Paragraphs: Effects on Apparent Theme, Reading Time, and Recall. *Journal of Verbal Learning and Verbal Behavior*, 17:13-28.

Lester, J. and Porter, B. 1997. Developing and empirically evaluating robust explanation generators: the Knight experiments. *Computational Linguistics*, 23(1):65-101.

Longacre, R.E. 1979. The paragraph as a grammatical unit. In *Syntax and Semantics: Discourse and Syntax*, Talmy Givon (ed), 12:115-134. Academic Press, New York.

Milosavljevic, M. 1999. *Maximising the Coherence of Descriptions via Comparison*, PhD thesis, Macquarie University.

Nunberg, G. 1990. *The linguistics of punctuation*. CSLI Lecture Notes, 18. CSLI Publications, Stanford, CA.

O'Brien, E., Raney, G.E., Albrecht, J.E. and Rayner, K. 1997. Processes involved in the resolution of explicit anaphors. *Discourse Processes*, 23(1):1-24.

Ostorm, J. and Cook, W. 1988. Better Paragraph Plus. Sixth Edition. HarperCollins, London.

Power, R. 2000. Planning texts by constraint satisfaction. In *The 18th International Conference in Computational Linguistics (COLING)*, Saarbrücken.

Sanford, T. and Lockhart, F. 1990. Description types and method of conjoining as factors influencing plural anaphora: a continuation study of focus. *Journal of Semantics*, 7:365-378.

Schriver, K.A. 1997. *Dynamics in Document Design: Creating Text for Readers*. Wiley Computer Publishing.

Stark, H.A. 1988. What do paragraph markings do? *Discourse Processes*, 11:275-303.

Van Dijk, T.A. 1997. *Text and context: exploration in the semantics and pragmatics of discourse*. Longman, London.

Walker, M., Joshi, A. and Prince, E. 1998. *Centering Theory in Discourse*. Clarendon Press, Oxford.

Wright, P. 1983. Manual Dexterity - a user-oriented approach to creating computer documentation. In *Proceedings of Computer Human Interaction*. pp. 11-18.