

# An Empirical Study of Multimedia Argumentation

Nancy Green

Department of Mathematical Sciences, University of North Carolina at Greensboro  
Greensboro, North Carolina 27402, USA  
nlgreen@uncg.edu

**Abstract.** We have analyzed a corpus of human-authored arguments expressed in text and information graphics, non-pictorial graphics such as bar graphs. The goal of our research is to enable intelligent argument generation systems to make effective use of these media. This paper presents and compares two classification schemes used to analyze the corpus, illustrated by examples from the corpus, and discusses implications for generation systems.

## 1 Introduction

In many domains of discourse, arguments employ quantitative data for support. Frequently, these arguments are expressed in a combination of text, tables of statistics, and information graphics such as bar graphs. The goal of our research is to enable intelligent argument generation systems to make effective use of text and information graphics. Towards this goal we are making an empirical study of the relationship of information graphics to text in a corpus of arguments. We have analyzed the corpus in terms of text coherence relations and argumentation strategies. In this paper, we survey some of the results of our work so far and discuss the implications for argument generation systems.

## 2 Related Work

In this section we describe related work in multimedia generation, focusing on issues relevant to computer generation of arguments in integrated text and information graphics.

Since the 1980's research in intelligent multimedia generation systems (IMPS) has addressed the problem of generating integrated text and graphics [1]. Early systems that generated text and pictorial graphics (illustrations, diagrams, and maps) identified several key issues in multimedia generation [2],[3],[4]. The media selection problem is the problem of selecting use of text or graphics for parts of the presentation. The simplest approach to media selection is for the choice to be built into the system, e.g., by encoding media-specific information into rhetorical strategies. A limitation of this approach is loss of flexibility and expressiveness. Another key issue, media coordination, is the problem of coordinating parts of a presentation expressed in different media. Early work in this area included generating multimodal referring expressions, references to things in the world made through a combination of natural language referring expressions and pictorial representations [2],[5]. These approaches required reasoning about the intended effects of components of the presentation in

different media. A related issue is the problem of detecting and resolving unintended effects [6].

Other IMPS research addressed the problem of generating natural language captions for computer-generated information graphics. Information graphics are non-pictorial graphics such as scatter plots and bar charts [7]. PostGraphe partially automated the generation of business reports [8]: the user could select a template indicating the main point of his desired report, e.g., to compare profits of one company to another; Postgraphe then would search among a set of graphic designs to create a graphic emphasizing the user's selected point, and would generate a caption based upon the selected design and template. Another system, the Caption Generation System [9] automatically generated captions for information graphics produced by a powerful automatic graphic design system. The captions were intended to help the user to interpret a graphic by describing complex aspects of its design and relationships between the graphic's elements and the user's data.

In contrast, the text generated by AutoBrief played a central rather than a supporting role in achieving the goals of a presentation [10],[11]. AutoBrief generated advisory presentations on transportation scheduling in integrated text and information graphics. In the AutoBrief architecture, the first stage of processing was to generate a media-independent plan for achieving presentation goals, without specifying how the acts of the plan would be realized in a particular medium. In the next stage, media selection, parts of the plan were assigned to the system's text generator and/or automatic graphic design component for realization. The AutoBrief architecture enabled the information graphics to be designed to achieve complex communicative goals. However, although some types of presentations generated by AutoBrief could be analyzed in terms of argument strategies, presentation plan structure in AutoBrief did not explicitly represent this type of knowledge. The need to consider argument strategies in media selection and media coordination was argued in [12]. Also, a later implementation of the AutoBrief architecture in a system for generating evaluative arguments explicitly represented evaluative argument strategies; the representation was used to influence several aspects of text generation such as lexical choice [13]. The goal of our corpus analysis is to gain knowledge that can inform the design of future generation systems such as AutoBrief. Previous corpus studies of multimedia documents [14],[15] focused on different properties than those that we studied.

### 3 Corpus

The focus of our corpus analysis is on use of information graphics in arguments intended for an educated reader who is not necessarily an expert in the domain of discourse. The present corpus consists of excerpts from the following :

- A report prepared by the U. S. Department of Energy on the topic of global warming, published as a hypertext document on the world wide web [16],
- Articles on bird census data from several issues of a newsletter for non-scientists, prepared by a research institution and published as a print quarterly [17],
- A college textbook on software process improvement for undergraduate computer science students [18].

Excerpts were selected that expressed an argument partly in text and partly through use of one or more information graphics or maps. We then analyzed the excerpts in two

ways: first, in terms of coherence relations, and second, in terms of higher-level argumentation strategies. Since the two approaches led to different insights, we shall discuss them separately.

## 4 Analysis of Coherence Relations

The goal of this analysis was to investigate the types of relations between text and graphics in our corpus as indicated by cross-media cue usage. We define a *cross-media cue* as a phrase used to signal explicitly a relationship between some text and a graphic in the same document, e.g., “Figure 9.4 shows that”, “(Figure 9.4)”, “(See Figure 9.4)”. Cross-media cues are similar in some respects to discourse cue phrases, connectives such as ‘although’ that help to convey discourse coherence relations, semantic and pragmatic relations between units in a coherent text. First, both help the reader to connect parts of a presentation, although in many cases the reader may be able to infer the connection. Second, for both types of cue, the same construction may be used to perform different functions, and the same function may be expressed in different ways. The expression of cross-media cues may involve use of discourse deixis, use of an expression that refers to the physical layout of a document, e.g., “the above Figure” [19]. Also, multimodal referring expressions can be used to indirectly signal a relationship between text and a graphic. However, we only considered explicit cues.

Rhetorical Structure Theory (RST) provides a set of coherence relations that can be used to describe the structure of a monomodal text [20]. According to RST, a common structural pattern is formed by two adjacent spans of text related by a coherence relation. One of the spans (the *nucleus*) is more essential to the writer’s purpose than the other (the *satellite*). We found that several RST relations could be used to describe the relationship between the graphic referred to in a cross-media cue and the rest of the sentence, provided that the graphic was analyzed as the nucleus of the relation and the rest of the sentence as the satellite. In the next section, we give examples of the relations that we found in our corpus, and the definition of the relation (as we adapted it to describe the multimedia discourse).

### 4.1 Coherence Relations

**Preparation.** Text prepares the reader to expect and interpret the graphic. Example: “Figure 1 on page 3 shows how average numbers of four common species have changed over the course of Project FeederWatch.”

**Restatement.** Text restates some or all of the situation expressed in the graphic. Example: “In the United States, nearly 85 percent of anthropogenic greenhouse gas emissions result from the burning of fossil fuels (Figure 3).” The figure contains a pie chart in which the slice denoting the contribution from the United States is labelled “85%”.

**Summary.** Text summarizes multiple pieces of information expressed in the graphic. There were two varieties of summary in the corpus. In the first example, the text makes a generalization about a subset of the data shown in the graphic, and in the second

example, the text expresses an arithmetic summarization of the raw data shown. Examples: “From Fig. 9.5, you can see that the numbers of test defects are much lower in the later programs.”, “Current projections show U.S. emissions increasing by 1.2 percent annually between 1995 and 2015 absent any policy interventions (see Figure 4).”

**Evaluation.** Text provides an evaluation of the situation conveyed by the graphic. Example: “In particular, anthropogenic carbon dioxide emissions have increased dramatically ... (Figure 1)”

**Elaboration.** Text provides additional information about the data not given in the graphic. Example: “These data, shown in Figs. 9.1 and 9.2, are for the times to both find and fix these defects.”

## 4.2 Discussion

This analysis shows a variety of relations between text and graphic that can be signaled by cross-media cue usage. However, it raises some issues for natural language generation. In general, what makes a multimedia document coherent? For example, are the relations between cross-media cue and graphic constrained by the coherence relations between spans of text in the document? Are there any relations needed to characterize multimedia discourse coherence that are not needed for monomodal text? We assume that these questions are relevant to argument generation, i.e., that in order to be effective, the presentation of an argument must be coherent. Another question is what factors determine where to place a cross-media cue? An overly simplistic approach might be to add a cross-media cue whenever a graphic stands in one of the above relations to the text. However, this approach is not sufficient when more than one unit in the text is related to the same graphic. In many cases, a graphic is relevant to more than one sentence. To address this problem, it is necessary to analyze the higher-level organization of the text, which is addressed in the next section.

## 5 Analysis of Argument Strategies

The goal of this analysis was to investigate the roles of text and graphics in argument strategies. We examined the corpus for instances of strategies described in a textbook on writing effective arguments [21]. The results can be summarized as follows:

1. A single graphic can be designed to convey more than one part of the same argument strategy, or to convey parts of several strategies.
2. The same part of an argument strategy may be expressed in both text and graphics.
3. In addition to bearing the argument proper, text in the body of a document may be what we shall call *commentary-bearing* rather than *argument-bearing*. The former category includes comment on
  - a graphic’s role in the argument, by means of cross-media cues;
  - the location of the graphic in the document, by means of discourse deixis;
  - correspondences between graphical elements and database elements, e.g., the type of data shown in a graphic; and

- salient visual features of the graphic, e.g., a sharp change in slope of a line in a line graph.
4. Placement of this commentary-bearing text may be interleaved with argument-bearing text. Furthermore, cross-media cues may occupy sites that would otherwise be available for discourse cue phrase placement.
  5. Intuitively, use of a graphic can add to some dimensions of a presentation's effectiveness, e.g., comprehensibility, memorability, or persuasiveness. (There is some support for these intuitions from cognitive psychology also.)

### 5.1 Argumentation Strategies

We now present detailed examples from the corpus to provide evidence for these observations. Excerpts are given in tabular format. The excerpt is shown in the order in which it appears in the source. The third column contains notes from our analysis. In the third column, *->* and *//>* denote *supports* and *refutes*, respectively; C denotes *commentary* and G the *graphic* referred to in the text. Each table is accompanied by a figure with a schematic of the design of any relevant graphics in the source.

**Addressing the Counterargument.** The text shown in Table 1 makes use of this strategy in units (5-7). Support is provided for a claim (5) by acknowledging (6) and refuting (7) a possible counterargument (not explicitly stated in the text; denoted by *ctr(5)* in the table). Both of the graphics in the source have the schematic form shown in Figure 1. They express the argument given in units (5-7); they show data supporting both (6) and (7), but (7) is emphasized over (6) by graphic design and layout. The design emphasizes Phase over Defect Type by encoding Phase by position on the x-axis for each cluster of bars and Defect Type by shading. The layout emphasizes Phase over Language by presenting data for each language in a separate graphic on different pages. Finally, note that a cross-media cue placement algorithm based only coherence relations between text and graphics would overgenerate; it would generate the same cue for each of (5), (6), and (7), since each provides a summary of data in the two graphics.

**Table 1.** Source: [18], p. 275

Unit	Excerpt	Analysis
1	<b>defect identification costs are highest during test and use.</b>	1 -> 2
2	<b>Thus anyone who seeks to reduce development cost or time should focus on preventing or removing defects before starting test.</b>	Main claim
3	<b>This conclusion is reinforced by the PSP data on the fix times for 664 C++ defects and 1377 Pascal defects ...</b>	C
4	<b>These data, shown in Figs. 9.1 and 9.2, are for the times to both find and fix these defects.</b>	C
5	<b>Fix times are clearly much longer during test and use than in the earlier phases.</b>	5 -> 2, [6,7] -> 5
6	<b>While this pattern varies somewhat between the two languages and by defect type,</b>	6 -> ctr(5), G -> 6
7	<b>the principal factor determining defect fix time is the phase in which the defect was found.</b>	7 //> ctr(5), G -> 7

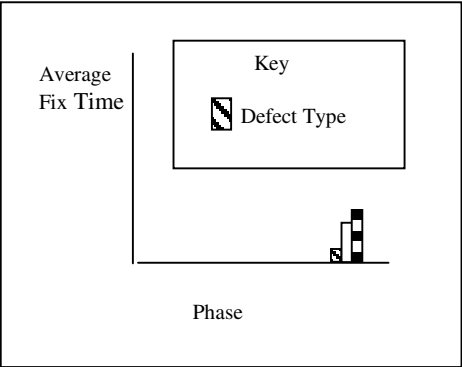


Fig. 1.

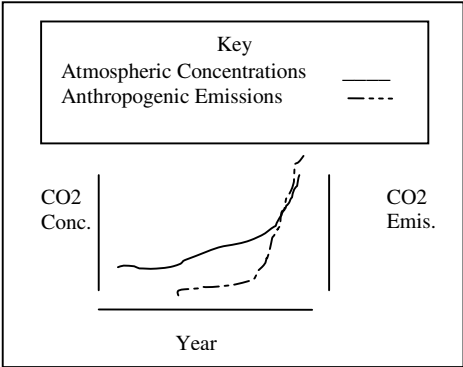


Fig. 2.

**Arguing for a Causal Relation by Showing Correlation (first example).** The text shown in Table 2 makes use of this strategy in units (1-6). The main claim, that something (4) caused a changing situation (1) during a certain time span (2), is supported by claiming that there was a proportional change in a related condition (5) over the same time span (6). The graphic has the schematic form shown in Figure 2. It shows data supporting both the claim in units (1-2) and units (5-6). In addition, the temporal correlation of (1) and (5) is shown by plotting the two sets of data against the same x-axis, which encodes time. The evaluation that anthropogenic emissions “increased dramatically” is reinforced by the steep rise in the line representing emissions.

Table 2. Source: [16]

	Excerpt	Analysis
1	Atmospheric concentrations of several important greenhouse gases (...) have increased by about 25 percent	G-> [1-2]
2	since large-scale industrialization began some 150 years ago.	Time span of 3
3	The growth in their concentrations is believed to be caused by	Main claim
4	human (anthropogenic) activity.	Main claim
5	In particular, anthropogenic carbon dioxide emissions have increased dramatically	[5-7] -> [3-4] G-> [5-6]
6	since the beginning of the industrial age	Time span of 5
7	due largely to the burning of fossil fuels and forestation (Figure 1).	Caused [5-6], C

**Arguing for a Causal Relation by Showing Correlation (second example).** The text shown in Table 3 makes use of this strategy in units (1-5), although it uses a graphic in a different way than shown in the preceding example. The main claim is a yes-answer to the rhetorical question (1). It is supported by claiming that a trend described in (2)

has features described in (4) that are correlated with events described in (5). (In addition, the trend in (2) is contrasted with the trend in (3) to show that winter weather affects northern wren populations differently than southern wren populations.) The graphic has the schematic form shown in Figure 3. It shows the trends in (2) and (3). The temporal correlation of features (4) of the trend in (2) with events in (5) is shown by annotating the peaks in the line denoting N.E. population with arrows and labels denoting the events (“snowstorms”).

Table 3. Source: [17], Autumn 1994, p. 5

	Excerpt	Analysis
1	<b>Did the bitter cold and frequent snowstorms this past winter take a similar toll on wrens in the Northeast?</b>	Main claim: yes
2	<b>We examined the weekly FeederWatch counts and noted that, in the Northeast Region, Carolina Wrens visited fewer and fewer feeders as the season progressed,</b>	G -> 2
3	<b>although they didn't follow this pattern in the Southeast region (Figure 7).</b>	G -> 3, C
4	<b>The Northeast decline came after a sharp peak in visitation during the week of January 8-14.</b>	G -> 4
5	<b>This period corresponded with a severe storm that dropped at least three feet of snow over most of the Northeast and glazed much of the Southeast with ice.</b>	Annotations in G

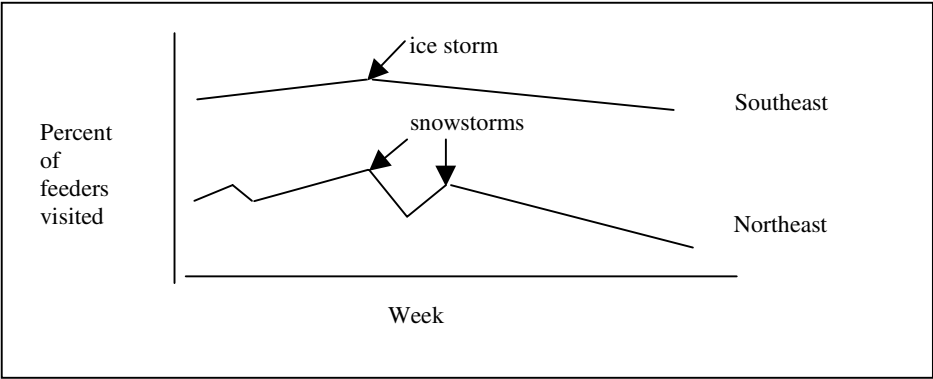


Fig. 3.

**Inductive Generalization.** The text shown in Table 4 makes use of this strategy more than once in units (1-5). A previous inductive generalization (1) was based on limited data described in (3). However, new data described in (4) falsify (1). Together, the old and new data support a new inductive generalization (2). Both graphics for the two whale species have the schematic form shown in Figure 4. The leftmost bar chart

shows data referred to in (3) supporting (1), while the rightmost bar chart shows data referred to in (4) falsifying (1); together the data in the two bar charts support (2). (5) repeats the argument in (3-4) for finback whales. In addition to the bar charts, the graphic contains a geographic representation of the parts of the Northern Hemisphere; arrows point to the region from where the data shown in the bar chart was collected.

Table 4. Source: [17], Summer 1995, p. 3

	Excerpt	Analysis
1	During the first year that BRP studied the world’s two largest whale species (blue and finback whales), our data suggested that they sang only at certain times.	Old claim
2	As it turns out, these species are vocally active all year round.	[3-5] -> 2
3	Initially, we monitored the region between Newfoundland and the Caribbean margin. There, we detected the most whale calls during the six-month-long winter breeding season.	G -> 3
4	More recently, however, we’ve recorded blue whales in the northeast Atlantic and we hear few or no calls in the winter there. Instead, the number of calls increases through late spring, peaks in late summer, and then decreases again in the fall – almost a mirror-image of the pattern farther to the south and west (Figure 1, top).	G -> 4. C
5	Finback sounds follow a similar pattern (Figure 1, bottom).	G -> 5, C

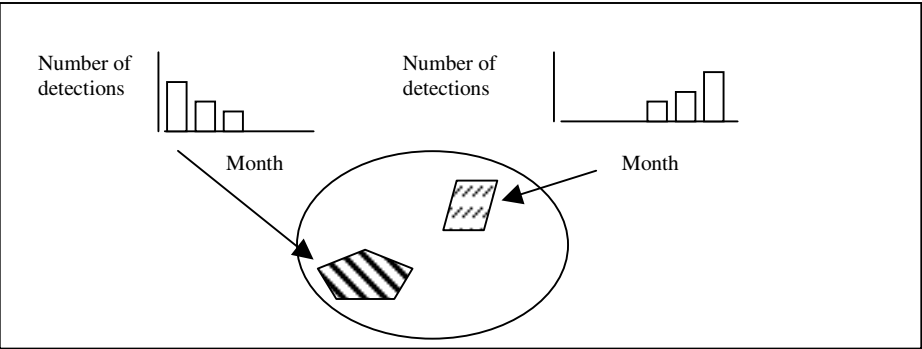


Fig. 4.

**Falsification of Hypothesis of Sufficient Cause.** The text shown in Table 5 makes use of this strategy in units (2-3) to argue that a condition (2) is not a sufficient cause of a situation described in (3) by claiming that the condition was present but that the situation did not occur. The graphic has the schematic form shown in Figure 5; the



bottom half of the figure, labeled *Figure 3b* shows data supporting the claim that the situation did not occur. Although the top half of the figure, labeled *Figure 3a*, is used with (4) as part of another argument strategy (addressing a possible counterargument), by juxtaposing it with Figure 3b the author enables the reader to see a striking visual contrast in the shape of the lines in the two graphics. Intuitively, seeing the two together helps to persuade the audience that (3) is an accurate interpretation of the data shown in Figure 3b.

Table 5. Source: [17]

	Excerpt	Analysis
1	... the “every other year” rule has its exceptions.	[2-3] -> 1
2	Even though conifers seems to produce abundant seeds only every second year,	
3	two conifer specialists, White-winged and Red crossbills, do not stage invasions in alternate years (see Figure 3b below).	G -> 3, C
4	Last winter was exceptional because both the regular biennial invaders (Figure 3a) and the less predictable crossbills traveled south together.	G -> 4. C

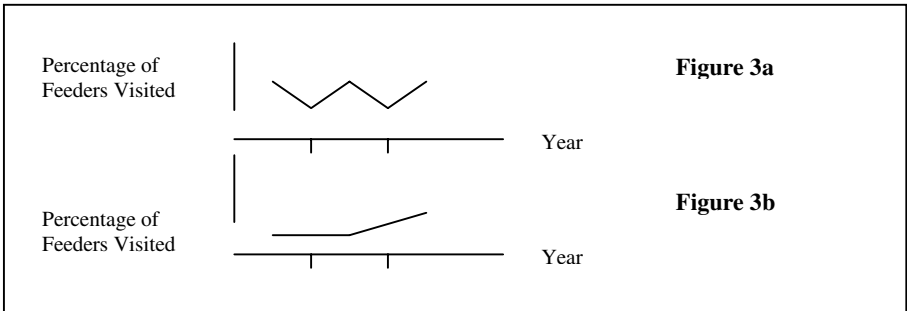


Fig. 5.

5.2 Discussion

This analysis has important implications for multimedia argument generation architectures. The influence of argument strategy can be seen in both text and graphics in the corpus. Thus, in a multimedia generation architecture such as described above, it would make sense to represent argument strategies in the media-independent plan explicitly. For example, this would enable media selection and graphic design processes to be informed by the underlying argument strategy. Second, we showed that the text serves two kinds of functions: to advance the argument proper and to help the reader to interpret a graphic and recognize its intended contribution to the argument. This suggests adding a post-graphic-design text-only planning phase to the architecture. This phase would be responsible for planning commentary-bearing acts (including cross-media cues) and integrating them with the already-planned argument-bearing acts previously allocated to text.

## Acknowledgments

This research was sponsored by a UNCG Summer 2000 Excellence Research Award.

## References

1. Roth, S.R., and Hefley, W.E. Intelligent Multimedia Presentation Systems: Research and Principles, in Maybury, M. T. (ed.). *Intelligent Multimedia Interfaces*. MIT Press, Boston, 1993, pp. 13-58.
2. McKeown, K.R., Feiner, S.K., Robin, J., Seligmann, D.D., and Tanenblatt, M. Generating Cross-References for Multimedia Explanation. *Proceedings of AAAI 1992*, 9-16.
3. Wahlster, W., André, E., Finkler, W., Profitlich, H.-J. P., and Rist, T. Plan-based integration of natural language and graphics generation. *Artificial Intelligence* 63(1993), 387-427.
4. Maybury, M. Planning multimedia explanations using communicative acts. *Proceedings of the Ninth National Conference on Artificial Intelligence*, 1991, 61-6.
5. André, E., and Rist, T. Referring to World Objects with Text and Pictures. *COLING-94*, 530-534.
6. Marks, J. and Reiter, E. Avoiding unwanted conversational implicatures in text and graphics. *Proceedings of the Eighth National Conference on Artificial Intelligence (AAAI-90)*, 450-6.
7. Card, S.K., Mackinlay, J., and Shneiderman, B. (eds.). *Readings in Information Visualization: Using Vision to Think*. Morgan-Kaufmann, 1999, ch. 1.
8. Fasciano, M., and Lapalme, G. Intentions in the coordinated generation of graphics and text from tabular data. *Knowledge and Information Systems*, Oct 1999.
9. Mittal, V., Moore, J., Carenini, G., and Roth, S. Describing Complex Charts in Natural Language: A Caption Generation System. *Computational Linguistics, Special issue on Natural Language Generation*. Vol. 24, issue 3, (1998), 431-467.
10. Green, N., Carenini, G., and Moore, J. A Principled Representation of Attributive Descriptions for Integrated Text and Information Graphics Presentations. In *Proceedings of the Ninth International Workshop on Natural Language Generation*, 1998, 18-27.
11. Kerpedjiev, S., and Roth, S.F. Mapping Communicative Goals into Conceptual Tasks to Generate Graphics in Discourse. *Proceedings of Intelligent User Interfaces (IUI '00), New Orleans, LA, Jan 2000*, pp 60-67.
12. Green, N. Some Layout Issues for Multimedia Argument Generation. *AAAI 1999 Fall Symposium on Using Layout for the Generation, Understanding or Retrieval of Documents*. Technical Report FS-99-04, 47-51.
13. Carenini, G. and Moore, J. A Strategy for Generating Evaluative Arguments. *Proceedings of the 1<sup>st</sup> International Conference on Natural Language Generation (INLG-00)*.
14. Green, N., Carenini, G., Kerpedjiev, S., Roth, S., and Moore, J. A Media-Independent Content Language for Integrated Text and Graphics. *Coling-ACL98 Workshop on Content Visualisation and Intermedia Representation*, J. Pustejovsky & M. Maybury eds.
15. Corio, M. and Lapalme, G. Integrated generation of graphics and text: a corpus study. *Coling-ACL98 Workshop on Content Visualisation and Intermedia Representation*, J. Pustejovsky & M. Maybry eds.
16. Energy Information Administration, U.S. Department of Energy. *Greenhouse Gases, Global Climate Change, and Energy*. Available at <http://www.eia.doe.gov/oiaf/1605/ggccebro/chapter1.html>.
17. Cornell Lab of Ornithology, Ithaca, New York. *Birdscope*.
18. Humphrey, W.S. *A Discipline for Software Engineering*. Addison-Wesley, 1995, pp. 274-283.
19. Paraboni, I., and van Deemter, K. Issues for the Generation of Document Deixis. In André et al. (Eds.), *Deixis, Demonstration and Deictic Belief in Multimedia Contexts, Proceedings of the Workshop associated with the 11<sup>th</sup> European Summer School in Logic, Language and Information (ESSLLI)*, Utrecht, The Netherlands, 1999, pp. 43-48.
20. Mann, W.C., and Thompson, S.A. Rhetorical Structure Theory: Toward a functional theory of text organization. *Text* 8(3) (1988), 243-281.
21. Mayberry, K.J., and Golden, R.E. *For Argument's Sake: A Guide to Writing Effective Arguments*. Second edition. HarperCollins College Publishers, New York, 1996.