

## The Construction of Novelty in Computer Science Papers

M<sup>a</sup> José Luzón Marco  
University of Zaragoza  
mjluzon@posta.unizar.es

### ABSTRACT

Novelty is a concept of great importance in the writing of a research paper, given that the author has to persuade the audience of the news value of the reported research, which makes it worth publishing. In this paper I use a corpus of computer science papers to investigate how novelty is created in this discipline. I analyse how the author uses evaluation and lexical cohesion to integrate his/her research into the existing knowledge structure of the field. Evaluation in computer science papers is closely associated with the Problem-Solution pattern which structures most of the papers: authors claim that the technology they introduce is the best solution to a problem that they have previously identified. Lexical cohesion highlights the novelty of the research by establishing a semantic relation of contrast between the fragment of text reporting previous research in the field and that reporting the authors' own research.

### 1. Novelty in scientific papers

Sociologists of science and researchers on scientific discourse have demonstrated that scientific knowledge is socially constructed by means of language and that scientific texts are not the objective report of findings, but rhetorical products whose aim is to persuade the readers (i.e. the scientific community) of the validity of the claims made there (see Knorr-Cetina, 1981; Latour and Woolgar, 1979; Yearley, 1981). An illustrative example of this view is Swales' (1990: 125) description of the research paper as a "key-product of the knowledge-manufacturing industry (...) cunningly engineered by rhetorical machining."

Scientists organise their accounts to represent their actions and beliefs in such a way that the research is presented as a new contribution to knowledge.

Bazerman states that in order to establish the value of a research within a particular market "the paper must first reconstruct the market, define the needs of the market and identify the research being reported as the proper vehicle for the satisfaction of those needs" (Bazerman, 1983: 169). This statement reflects the structure of Introductions as defined by genre analysis. Swales (1990) states that introductions of research papers tend to have three moves. The first one is *Establishing a territory* ("reconstruction of the market"), where the author draws attention to a topic (by claiming centrality, making topic generalizations, and/or reviewing items of previous research) and indicates his/her commitment to the research done in this field. The second move is *Establishing a niche* ("definition of the needs of the market"), where the author claims that the existing knowledge in a particular field must be modified in some way (usually by asserting that it is inadequate and should be increased). This move, which usually provides a negative evaluation of previous research (see section 2.2), is realised by counter-claiming, by indicating a gap or inconsistency in previous knowledge, by raising questions and pinpointing problems or by continuing and adding to a tradition. The third move is *Occupying a niche* ("identification of the reported research as the means to satisfy the needs"), where the author makes a proposal for introducing a new piece of knowledge, which involves the modification required in the previous move. In order to establish credibility and to show the originality and value of the research reported, the author should, therefore, show that he/she is familiar with the field, identify some needs in the field, and persuade the audience that the research provides a novel and appropriate solution.

Several authors (Bazerman, 1985; Huckin, 1987; Kaufer and Geisler, 1989) have shown the importance of conveying the idea that the research reported in an academic paper is at the "cutting edge", that it is a novel contribution to the knowledge in the field. Researchers on novelty usually emphasise the need to reach a compromise between two opposing rhetorical demands or two competing impulses when writing a research paper (Kaufer and Geisler, 1989; Myers, 1990): the impulse to present information inherited from a cultural community and the impulse to report something new. When defining the cognitive and social constituents important for a general theory of novelty, Kaufer *et al.* (1989: 290) mention the following four aspects: (i) the author should represent prior consensual knowledge, since the claim for newness is a claim against some aspect of this knowledge; (ii) the author should put forward systematic claims against consensual knowledge, pinpointing gaps or inconsistencies in this knowledge; (iii) the author must develop an argument to support novelty claims; (iv) and the author expects to achieve credit for the new knowledge. In order to be worth publishing the piece of knowledge reported must be not only original but also relevant to the existing research and knowledge in the field. And this can only be done if the work is placed in the context of existing research in the field and examined in the light of previous research. As Berkentotter and Huckin (1995: 47) point out: "It is only when scientists place their laboratory findings within a framework of

accepted knowledge that a claim to have made a scientific discovery and thereby to have contributed to the field's body of knowledge, can be made."

Myers (1990: 67) remarks that papers require a negotiation between the writer and the audience: while the researcher tries to show that he/she deserves credit for a new piece of knowledge, the editors and reviewers try to relate the claim to previous knowledge relevant for the community. The published paper has been defined as "a multilayered hybrid co-produced by the authors and by the members of the audience to which it is directed" (Knorr-Cetina, 1981: 106). Berkentotter and Huckin (1995: 39-41) point out, quite rightly, that the need to integrate credibility and novelty is reflected in the kind of rhetorical moves in specific sections of papers: the Introduction and the Discussion. The Introduction usually offers a new solution to a problem, and shows the author's familiarity with previous research intended to solve the problem; the Discussion always contains the author's claim to having occupied the niche in the Introduction, and it usually includes a comparison of the author's research with previous research or findings, with the indication that the results deriving from the current research are different or superior. Section 2 in this paper shows how the different moves contribute to the construction of novelty in computer science papers.

The aim of this paper is to reveal how novelty is constructed in papers in the field of computer science and to show the role of evaluation and cohesion in orchestrating for novelty. Although novelty has been investigated in academic papers in general and in scientific disciplines such as Medicine or Biology (e.g. Berkentotter and Huckin, 1995; Kaufer and Geisler, 1989), it has been devoted less attention in the area of technology. Research in computer science seems to have specific features which will determine the conventions of computer science papers. Cooper (1985) found out that, as opposed to the high use of referencing in other fields, the citation of literature was not so common in her corpus of IEEE publications dealing with advances in computer technology. She suggests that this fact is related to the features of the discourse community and the genre in this field, such as the little accumulated research tradition in the field, or the fact that the research is usually product-related. According to Swales, Cooper's data suggest that "evolving discourse communities on the periphery of the academic world may be developing alternative conventions for their central genres" (Swales, 1990: 148).

## **2. The study**

### **2.1. Corpus and Method**

A corpus of 30 papers was selected from ten journals in the field of computer science (see Appendix). Although references to previous work and claims of novelty may occur throughout the paper, I have focused on the introductory sections and on the Discussion/Conclusion, the sections where these rhetorical elements are most likely to occur (Swales, 1990; Berkentotter and Huckin, 1995).

I have tried to identify the linguistic/rhetorical strategies used by the author to integrate his/her research into the existing knowledge structure of the field and, in this way, construct novelty. Swales' model of paper introductions has been taken as a major reference, since it shows how successful writers of academic papers construct novelty, positioning their work within the field.

For the analysis of the role of evaluation in the construction of novelty I bear in mind the view of evaluation as a discourse product proposed by Hunston (1993, 1994) and Thetela (1997). According to these authors, evaluation is not exclusively associated with attitudinal language, but occurs whenever there is an indication of the writer's attitude to the value of any information presented in the text. Evaluation is created in discourse and depends on the value system underlying the genre and the goals of the text writer. Given that features such as originality, relevance or certainty are highly valued in the research paper, the writer must show that his/her research has these features. Thetela (1997) considers that there are two main concepts to analyse evaluation: "ascribed value" (AV) (i.e. "the evaluation bestowed on a thing") and "evaluated entity" (EE) (i.e. "the thing in the real world on which value is bestowed"). I will, therefore, examine which elements are evaluated and which values are ascribed to them.

For the analysis of the role of cohesion I draw on Hoey's work on lexical cohesion (1991), which shows that cohesion resulting from lexical relations contributes highly to coherence and that the interpretation of the relation between (adjacent or non-adjacent) sentences linked by a certain number of cohesive ties gives way to a form of text organisation. Hoey shows that there is a semantic relation between the sentences that are linked by lexical cohesive relations, which makes it possible for us to read them in the context as an intelligible pair. Hoey considers sentences that are linked in the light of the semantic relation between sentences (Winter, 1977) and of the patterns of organisation (Hoey, 1983) which involve the semantic relations between larger stretches of text. Lexical cohesion is considered here a pragmatic process (see McCarthy, 1990), which means that the items that collocate must be analysed in relation not to their place in an abstract system of meaning but to their place in the actually occurring text.

## 2.2. Evaluation

As was pointed out in section 1, in order to show the originality of the research, the author needs to place the research in the context of existing knowledge in the field, which is done in the move "Establishing a territory." The first step in this move ("Claiming centrality") consists in a positive evaluation of the research area or of an aspect of the field, in order to get the discourse community to accept the importance of doing research in this area. In our corpus this positive evaluation was effected in one of the following ways:

(i) with words which point to the growth of *interest* in the area or in the use of a technique or approach:

- (1) a. Medical *use* of computer-generated 3D images to display anatomic data *has grown enormously in recent years*. (Segall et al.)
- b. The use of real-time embedded systems (RTESs) is becoming *more and more widespread* in *today's* computing/electronic climate. (Fox et al.).
- c. There is *mounting* evidence that asynchronous circuits are finding a niche in high-performance applications. (Chakraborty et al.)
- d. *Interest* in the use of triangulated models has *increased recently*... (Jones et al.).

There tend to be time references that highlight the up-to-date status of the research area (e.g. "in recent years", "the last decade", "today's", "recently"). This is also emphasised by the use of the present perfect and the present progressive.

There may be references to the high number of people working in the area, suggesting, therefore, the interest of the research for a great part of the community:

- (2) a. Many have found it interesting and valuable to demonstrate light interference in computer-generated imagery (Sun et al.).
- b. Analysing flow-like patterns has been *an active* research topic *for certainly more than one decade*; see for instance Kass and Witkin (Kass and Witkin, 1987) (Weickert).

(ii) with items which refer to the intrinsic *importance* of the area or aspect under research:

- (3) a. To help the embedded-system designer meet the requirements of modern systems, tool support is a *key* issue. (Fox et al.)
- b. The number of image database systems which have recently been developed is *strong evidence to this area's importance*. A *key* aspect of image databases is the creation of robust and efficient indices. (Androutsos et al.)
- c. Decomposition is *central* to methodologies based on a stepwise refinement of system analysis and design (Go et al.).

*Interest* and *importance* are positive values, ascribed to the area or field within which the reported research is developed, and which add to the value of such research.

The move "Establishing a territory" can also be realised by making topic generalisations and/or by reviewing, and usually evaluating, previous research (Swales, 1990). In computer science papers the writer does not usually report on previous findings concerned with knowledge per se, but on previous technical solutions to an established problem (previous models, tools, etc.). Since the review of previous research usually involves evaluation of the reviewed items, such review links Move 1 ("Establishing a territory") and Move 2 ("Establishing a niche"). However, there are some cases where the reference to previous work does not involve evaluation.

In some cases authors refer to previous work to lend support to a general statement and to reveal where further information about the topic can be found:

- (4) a. A hand-held probe provides poor positional accuracy- unlike scanning, which produces accurate, continuous data\$1 (Segall et al.).
- b. Content-based image retrieval (CBIR) is a research area dedicated to the image database problem [1] (Androutsos et al.).

or to provide a general framework and show familiarity with the area and with the work of other researchers in such area:

- (5) a. Pioneering simulations of this phenomenon include interference fringes by Smits and Meyer [5] (...) Recently Schramm et al. (6) presented a detailed analysis of spectra and structures (Sun et al.).
- b. Examples of net driven techniques are the elimination of immediate transition in GSPN (12), the exploitation of symmetries in Stochastic Well-Formed Colored Nets (13) (Campos et al.).

References can also be used to support centrality claims:

- (6) a. Interest in the use of triangulated models *has increased recently* (DeLucía and Black 1987; Jones et al. 1995) (Jones et al.).
- b. The subject of combining the finite element structural analysis technique and numerical optimization algorithms *has been widely investigated* in the last two decades (1) (Song et al.).
- c. Analysing flow-like patterns has been *an active* research topic *for certainly more than one decade*; see for instance Kass and Witkin (Kass and Witkin, 1987) (Weickert).

As has been said, the reference to previous work usually involves evaluation. This is usually a negative evaluation, necessary to show that the reported research is a new contribution to knowledge.

However, the writers can also evaluate previous work positively when the results contribute to showing the importance or usefulness of their own work:

- (7) Kriegal et al. (1991) *show* that considerable benefits in spatial query can be obtained at the refinement step by employing structural decomposition of the spatial objects (Jones et al.).

or the appropriateness of their method:

- (8) This (method) *has previously been shown* to provide good performance in proximity searches (Hoel and Samet 1991, 1992; Hjaltason and Samet 1995) (Jones et al.).

*Show* is a reporting verb that evaluates positively the following information, bestowing it the status of certain (Thompson et al., 1991).

The writers may refer to others' work or they may refer to their own previous research. They cite themselves to establish their authority in a specific area of research, to draw attention to their own previous accomplishments in the field and to show that this new piece of research is a continuation of their line of research. In the following example reference "[9]" is a reference to the authors' previous work, on which the present paper is based. Their own previous work is evaluated positively, since it is stated that it was a step to solve the problem posed in the paper:

- (9) The main problem is that (...) the product space may contain nonreachable states (...) Two level techniques *proved very effective* in a number of cases. For certain net subclasses like MG no spurious states are generated so that equality between RS and RSP is obtained (6). In [9] the MG case is *generalized* to DSSP where RPS may strictly include RS.

In this paper the extension of the two level approach for DSSP is developed for arbitrary but bounded SPN models (...) *The paper is based on a previous work* [9] for the restricted DSSP class. A *significant improvement* is that the method presented in this paper applies to arbitrary SPN systems (Campos et al.).

The second move in Swales' model, "Establishing a niche", functions as a link between Move 1 ("Establishing a territory") and Move 3 ("Occupying the niche"), between the reference to previous work and the introduction of the author's work. The move "Establishing a niche" usually consists of a negative evaluation of previous research, which justifies the introduction of the authors' own research. Swales states that there are several ways to establish a niche in research papers (Counter-claiming, Indicating a gap, Question-raising, Continuing a tradition). Using a corpus of computer science papers, Posteguillo (1999: 143) found that the preferred means of presenting the need for the work was indicating a gap. Our research confirms this finding and reveals that the step "Indicating a gap" tends to be realised in a particular way in computer science papers. The gap does not usually refer to a gap in knowledge but to a gap in technological development. That is, rather than a knowledge-related gap, it is a product-related gap, as the following example illustrates:

- (10) Specifically the color histogram remains the most popular index, due primarily to its simplicity [6,7]. However, using the color histogram for indexing has a number of *drawbacks* (Androutsos et al.).

In the example above the group "a number of drawbacks" is a cataphoric label (Francis, 1994): an unspecific nominal element which is realised lexically in the subsequent text. This nominal group frames the subsequent discourse in a negative evaluation. What follows is a list of the "drawbacks" of color histograms:

- (10a) Histograms *require* quantization (...) However, with quantization comes *loss of color* (...) *Color exclusion* is difficult using histogram techniques (...) Histograms can provide *erroneous retrieval results* (...) *This poses a problem*. For image retrieval this can cause *very poor performance* (...) The histogram captures global color activity; *no spatial information is available*.

Negative evaluation of previous work does not only occur in Move 2 ("Establishing a niche") of the Introduction. Some papers include, after the Introduction, sections where previous research related to the specific topic of the paper is described and evaluated quite extensively (e.g. sections with the title "Related Work"). These sections present previous solutions to the problem posed in the Introduction (previous methods, tools, models), which are usually dismissed as inadequate.

- (11) Most wormhole routing schemes based on avoiding deadlock *limit the adaptivity* of the routing algorithm. Oblivious deadlock avoidance schemes *do not allow* any routing adaptivity. This *curtails performance* by *preventing packets from having the ability* to route around congestion and faults. Other avoidance schemes allow fully adaptive routing, but require *an increased number of virtual channels*. (...) Excess virtual channels can *increase router complexity*, leading to *longer router clock cycle times* and a consequent *performance degradation*. (...) Hence, schemes based on deadlock avoidance generally *suffer from losses in adaptivity and increased router complexity*, both of which can *negatively impact* performance (Pinkston).

The third move in the Introduction, "Occupying the niche", is characterised by Swales (1990: 159) in the following way:

The role of Move 3 is to turn the niche established in Move 2 into the research space that justifies the present article. The link between the moves is a strong one. When a Move 2 occurs (...) the ensuing Move 3 variously offers to substantiate the particular counter-claim that has been made, fill the created gap, answer the specific question or continue the rhetorically established tradition.

Thus, while Move 2 involves a negative evaluation of previous research or technology, Move 3 is characterised by the use of positive evaluation. The following fragment illustrates the close relation between Move 2 and Move 3.

- (12) MOVE 2

Removing a lenticular volume of tissue with a central thickness of approximately 12 microns cause one diopter of ocular refractive change. **However**, *current* available technology *lags behind* in the ability to monitor the anatomical changes that occur as a result of the procedure.

*Current* optical systems for measuring corneal and epithelial thickness (...) *suffer from a significantly lower level of precision* (...) and *won't work* if the corneal



transparency is compromised. Furthermore, they *don't permit* accurate topographic localization of the measurement point and therefore have *limited use* in analyzing thickness profiles.

Currently available hand-held ultrasonic pachymeters can measure corneal thickness through opaque tissues, **but** use transducers operating at frequencies up to 20 MHz (at this frequency the corneal epithelium *cannot be resolved*) (...) Furthermore, a hand-held probe provides *poor positional accuracy*, **unlike** scanning, which produces accurate, continuous data (...)

### MOVE 3

In this article we describe a software application developed in ANSI (...) We designed the application to analyze and display positional data derived from our VHF ultrasound scanning system for surfaces of the epithelium, stroma, and other pathological layers (...) It is the first system to *enable examining individual corneal layer thickness and surface shape simultaneously* (Segall et al.).

The first three paragraphs in the fragment (Move 2) are concerned with the gap in technology. The available technology is evaluated negatively (e.g. "lags behind", "suffer from a lower level of precision", "won't work"), to show that there is a gap that needs to be filled. There are time adjectives and adverbs which indicate that the applications being evaluated negatively are the current ones. The available technology is evaluated in terms of performance, by highlighting its limitations. The last paragraph (Move 3) is concerned with the technology the authors propose, which is evaluated positively, in terms of performance ("enable examining...") and of novelty ("It is the first system to...").

As example (12) illustrates, in our corpus the move "Establishing the niche" usually takes the form of presentation of a problem. The introductions in the papers analysed usually have a Problem-Solution pattern, where the move "Establishing the niche" defines the problem and the move "Filling the niche" identifies the solution.

Discourse patterns, such as the Problem-Solution pattern, the Hypothetical-Real pattern, and the Unexplained-Explanation pattern, have been used to characterise the structure of research papers. The papers produced by a specific discourse community may be structured following specific patterns. Adams Smith (1987, 1990) states that biomedical research papers should be explained in terms of the Unexplained-Explanation pattern, consisting of the following elements: Situation-Unexplained-Procedural-Findings-Interpretation-(Evaluation)-Explanation. The Hypothetical-Real pattern seems to be highly appropriate to write economics texts, since economics writers need to move between the ideal world of models and the world of reality (Hewings, 1990). The Problem-Solution pattern (Hoey, 1983), which provides a framework for most computer science papers, consists of two basic parts narrowly related: *problem* and *solution*. They may be preceded by a *situation* part, which establishes the basis for the problem, and followed by an *evaluation* part, which provides the assessment of the solution. The following words by Berkentotter and Huckin (1995: 33) help to explain the frequent use of the Problem-Solution pattern in computer

science papers: "We would expect the journals to gradually modify their genre conventions so as to better accommodate the needs of specialist readers and readers pressed for time (who may, in many cases, be one and the same). " Readers of computer journals are mostly concerned with finding tools that solve the problems that this fast developing area of research meets. Usually in the Introduction of the paper the problem is presented and the solution is put forward. In the Discussion the solution is reasserted and evaluated (or the evaluation already effected in the Introduction is developed further or justified).

The pattern we find in most computer papers may be represented as follows: Situation (statements about the field of research)- Problem-Solution(s) (others' solutions)- Negative Evaluation of the Solution(s)- New Solution (the authors' solution)- Positive Evaluation. After stating a problem or need, the solutions that have been proposed previously are described and evaluated negatively. Finally, the solution given in the paper is presented as the best one and the one that solves the problem that the other pieces of research or the current technology cannot solve. The following example illustrates this pattern clearly:

#### (13) SITUATION

To increase communication efficiency, many multicomputer interconnection networks incorporate wormhole switching (1), virtual channel flow control (2) and adaptive routing techniques (3). Respectively these techniques pipeline, multiplex and increase flexibility in the use of networks bandwidth, allowing routing to be performed more efficiently.

#### PROBLEM

However, in implementing these techniques, the critical issue of deadlocks *must be addressed* (...) *Progress in routing halts* when packets block waiting for channel resources (physical or virtual) while holding onto other channel resources. Knotted cyclic dependencies formed by groups of such blocked messages lead to the indefinite *postponement of messaging activities* and, if allowed to persist, could result in the *entire network coming to a complete standstill*.

#### SOLUTION AND NEGATIVE EVALUATION

Avoidance has been the traditional approach for addressing deadlocks in routing (...) Routing schemes which restrict routing options in order to avoid deadlocks *are neither maximally flexible nor efficient*. An alternative approach for addressing deadlocks in routing is deadlock recovery (...)

#### SOLUTION AND POSITIVE EVALUATION

This paper presents the notion of progressive deadlock recovery routing, which *enables the design of flexible and efficient routing algorithms capable of handling deadlocks effectively*. This approach *allows* true fully adaptive routing of packet in which unrestricted routing on all virtual channel over all router ports is permissible (...) This approach offers the following significant advantages (Pinkston).

In some cases the authors include among the previous solutions their own previous solution, and the solution given in the paper is presented as an improvement.

The Problem Solution pattern which structures the paper is also very frequently used to organise the abstract:

- (14) (SITUATION) Color is the characteristic which is most used for image indexing and retrieval. Due to its simplicity, the color histogram remains the most commonly used method for this task. (PROBLEM) *However, the lack of good perceptual histogram similarity measures, the global color content of histograms, and the erroneous retrieval results due to gamma nonlinearity, call for improved methods.* (SOLUTION) We present a *new* scheme which implements a recursive HSV-space segmentation technique to identify perceptually prominent color areas. (...) (EVALUATION) Our system provides *accurate retrieval results and high retrieval rate*. It *allows for* queries based on single or multiple color and, in addition, it *allows for* certain colors to be excluded in the query (Androutsos et al.).

At the end of the Introduction or at the beginning of the Conclusion, where the author introduces the main claim of his/her research or reasserts this claim, the paper often includes explicit references to the novelty and originality of the paper (such as the adjectives *new, novel, first*, or explicit references to the contributions of the paper). This reflects the underlying ideology in the process of writing a scientific paper: there is a need to contribute something new to the already existing knowledge.

- (15) a. It is the *first* system to enable examining individual corneal layer thickness and surface shape simultaneously (Segall et al.).  
 b. Section 3 presents recently proposed deadlock recovery routing techniques and introduces a *new* progressive recovery routing approach. (Pinkston)  
 c. Although the proposed algorithm is based on an idea similar to Myers and Meng's algorithm (15), there are several important *extensions to* and differences from their work. We believe that these represent *important contributions to* the existing body of work on... (Chakraborty et al.).

These explicit references to novelty may also occur in the title, as the following two titles illustrate:

- (16) a. A *novel* vector-based approach to color image retrieval using a vector angular-based distance measure. (Androutsos et al.)  
 b. A *novel* node-based structural shape optimization algorithm (Song et al.).

Stating the novelty of the technology introduced implies a comparison with previous technology or with previous research. In the Conclusion this comparison is frequently explicit, thus highlighting the news value of the research:

- (17) This *new* method for transforming colors into spectral is heuristic, *but more accurate and flexible* than previous transformation methods (Sun et al.).

The main function of the Conclusion is to "fill the gap" of the Introduction. It serves to reintroduce the claim for novelty after this claim has been substantiated in the preceding sections. Berkentotter and Huckin (1995: 47) point out that the Discussion has the same types of moves as the Introduction, *but in reverse order*:

In Swales' terms, the writer asserts his or her claim to the niche created in the Introduction and occupied in the subsequent investigation. Second, there is a series of statements showing how these results respond to the larger issue stated in the Introduction. These statements usually include a comparison of the present results to related findings or claims made by other investigators

In the papers analysed the Conclusion tends to be used by the authors to *emphasise the novelty of the technology they present*. The first sentence is frequently a presentation of the technology they have introduced in the paper. This is followed by a novelty claim, where the contribution of the research to the field is stressed.

- (18) We have presented here a *new* method for decomposing process specifications (...) The method overcomes the three restrictions of our earlier method (14) It can decompose general processes, including recursion, nondeterminism and internal action (Go et al.).

### 2.3. Cohesion

The end of the Introduction and the beginning of the Conclusion indicate that the paper, or the technology proposed in the paper, solves the problem posed previously in the Introduction. This is done by means of cohesive devices and coherence relations. Cohesion does not depend on formal relations *but on the reader's knowledge of the underlying semantic structure of the text*. This is specially true of scientific papers, where, as Myers (1991: 5) points out, "readers must have a knowledge of lexical relations to see the implicit cohesion."

Cohesive links play an important role in the construction of novelty, since the author uses them to establish a contrast between previous technology or research and his/her own research. The role of cohesion in the construction of novelty is illustrated in example (12) above. Lexical cohesion is a *pragmatic process* (see McCarthy, 1990). This implies *analysing the items that collocate in terms of their function in the actually occurring text*. The first three paragraphs in example (12) describe current available technology. The entity "*current available technology*" is specified later as "*current optical systems for...*" and "*current available hand-held...*" The strings of words "*lags behind*", "*suffer from a lower*

level of precision", "won't work", "don't permit", "limited use", "cannot be resolved" and "poor positional accuracy" collocate because they all express negative evaluation.

The cohesive relations between the two parts of the fragment (Move 2 and Move 3) make it possible for us to interpret these parts as related elements which form a pattern. To analyse how cohesion contributes to novelty we have to bear in mind Hoey's finding that the interpretation of the relation between sentences linked by a certain number of cohesive ties gives way to a form of text organisation. The construction of novelty is effected through a *Matching relation*, whose underlying semantics is *comparative affirmation* and *comparative denial* (Winter, 1977: 30). The typical matrix clause of comparative denial is "What is true of X is not true of Y." In example (12) the contrast between current technology and the technology the authors introduce begins at the end of the third paragraph and is signalled by the preposition "unlike": the "poor positional accuracy" of a hand-held probe is contrasted with the "accurate data" produced by scanning. The preposition "unlike" marks a break between different discourse segments, which should be recognised by the reader to understand the text. At the end of the fourth paragraph "enable" signals a positive evaluation, in this way establishing a contrast with all the previous negative items.

The following example illustrates a common strategy used by writers of computer science papers: to present a list of problems or drawbacks of current technology and to state in another part of the paper (end of Introduction or Conclusion) the advantages of the technology that they propose, which contrast with the previous drawbacks. The example has been taken from a paper with the title "Flexible and efficient routing based on progressive deadlock recovery." As can be seen, the title is a paraphrase of the statement where the authors introduce their claim (First line in part A). Part A occurs at the end of the Introduction, while part B occurs at the beginning of the section "Related work."

#### Part A

This paper presents the notion of *progressive deadlock recovery routing*, which enables the design of flexible and efficient routing algorithms capable of handling deadlocks effectively. This approach *allows true fully adaptive routing* of packet in which unrestricted routing on all virtual channel over all router ports is permissible (...) This approach offers the following significant advantages:

1. *true fully adaptive routing.*
2. *no virtual channels required for handling deadlocks,*
3. *simple design of fast and flexible routing*
4. *predictable and efficient recovery from deadlock* (Pinkston)

#### Part B

*Most wormhole routing schemes based on avoiding deadlock limit the adaptivity of the routing algorithm.* Oblivious deadlock avoidance schemes *do not allow any routing adaptivity.* This curtails performance by preventing packets from having the ability to route around congestion and faults. Partially adaptive avoidance schemes allow some

amount of adaptivity, but *not full adaptivity*. Other avoidance schemes allow fully adaptive routing, but *require an increased number of virtual channels*. (...) Excess virtual channels can increase router complexity, leading to *longer router clock cycle times* and a consequent *performance degradation*. Moreover, virtual channels devoted to deadlock avoidance can *underutilize network link bandwidth*- bandwidth that might *otherwise be used more efficiently* with unrestricted use if deadlocks occur rarely. Hence, schemes based on deadlock avoidance generally suffer from losses in adaptivity and increased router complexity, both of which can negatively impact performance (Pinkston)

"Progressive deadlock recovery routing" (part A) is opposed to "schemes based on avoiding deadlock" (part B). The advantages of the new systems are presented as overcoming the drawbacks of the current systems: (1) "true fully adaptive routing" vs. "limit the adaptivity of the routing algorithm", "do not allow any routing adaptivity", "not full adaptivity"; (2) "No virtual channels required for handling deadlocks" vs. "require an increased number of virtual channels"; (3) "Simple design of fast and flexible routing" vs. "longer router clock cycle times"; (4) "predictable and efficient recovery from deadlock" vs. "performance degradation", "underutilize network link bandwidth."

The identification of the pairs of sentences that are related must be accounted for both in terms of repetition links and coherence processes (Hoey, 1991). The latter include processes such as lexical equivalence, expansion, reduction or transference. For instance, when establishing a contrast relation between "predictable and efficient recovery from deadlock" and "underutilize network link bandwidth" the writer assumes that the reader knows what "efficient" implies in this context.

### 3. Conclusions

Novelty is an important element in the current trend of promotion of scientific research. Berkentotter and Huckin (1995: 43) suggest that there are two obvious reasons for this need to promote one's own work: the increased competitiveness of modern science, and the fact that modern science is embedded within modern "promotional" or "consumer" culture. The scientific journal is the instrument used by scientists to promote their work, which makes it evolve towards promotional genres. This seems to be specially true of computer science papers, used by the authors to introduce new models and new technologies and to show the advantages of this technology over that currently available.

This paper has revealed the devices used by writers of computer science papers to construct novelty, to show that their work is a new and valuable contribution to the field which solves the problems that previous work could not solve. The authors use evaluation and lexical cohesion to create a semantic relation of contrast. Previous research is evaluated negatively in terms of performance: it is inadequate and it does not satisfy the requirements necessary to meet the need identified in the paper. The authors' own research is evaluated



positively in terms of three parameters: interest or importance (the area within which the research is placed is an important one or one that arises a great deal of interest at the moment of research), performance (it has the required features to meet the need identified in the paper), and originality (no previous work or research has met this need so fully). The use of evaluation is closely associated with the Problem-Solution pattern that structures most computer science papers. These papers are usually structured with the following pattern: Situation (interest of the field of research)- Problem- Solution(s) (others' solutions)- Negative Evaluation of the Solution(s)- New Solution (the authors' solution)- Positive Evaluation.

The use of lexical cohesion to create novelty cannot be separated from evaluation. The items used to offer a negative evaluation of previous research collocate because they are part of the same schema. Similarly, the items used to offer a positive evaluation of current research collocate because they all contribute to the positive presentation of the research. The cohesive links between these two sets of items give way to a pattern of organisation, which makes it possible to read the two fragments where the positive items and the negative items occur as an intelligible pair. Usually there is a correspondence between positive and negative items, artfully created by the authors to show that the current research/technology overcomes all the deficiencies or drawbacks of previous research/technology.

This research has pedagogical implications for the writing of computer science papers. The first one is the importance of teaching the values that underlie research papers and the rhetorical function of evaluation in these papers, and of showing how evaluation is effected. The second implication is the need to make the students aware that language is patterned and that computer science papers tend to follow the Problem-Solution pattern. Cohesion, and more particularly lexical cohesion, should be presented as a device useful for the creation of patterns. Finally, it is important to help the students realise that the devices used to create novelty in computer science papers depend on the constraints of the genre in this discipline and on the features of the discourse community.

Given that the scope of our research has been necessarily restricted, it seems appropriate to make caveats. This study has been limited to the analysis of novelty in computer science papers. Therefore, the findings should not be extrapolated for papers in other disciplines. Berkentotter and Huckin (1995) have shown that scientific papers have evolved to incorporate or develop features which help carry news value (e.g. an informative title, a results-oriented abstract, powerful visual aids, prominent statements of findings and claims). And, as we have seen in section 1, the rhetorical structure of the introduction of scientific papers is intended to foreground novelty. However, as Cooper (1985) claims, computer science papers have specific features deriving from the constraints of the discipline. It would be interesting to investigate whether papers in other disciplines use the same mechanisms to construct novelty. This analysis might reveal that novelty is more important in some disciplines than in others. The analysis of novelty in humanistic papers could also produce significant results. Since the discourse of scientific and humanistic disciplines is related to different communities and social contexts, it can be expected that the authors will make use of different discourse strategies.

## Works Cited

- Adams Smith, Diane (1987): "Variation in field-related genre." *ELR Journal* 1: 10-32.
- (1990): "Source and derived discourse." In M.A.K. Halliday, J. Gibbons, and H. Nicholas, eds., *Learning, Keeping and Using Language. Vol. II. Selected Papers from the 8th World Congress of Applied Linguistics*. Amsterdam: John Benjamins, 415-433.
- Bazerman, Charles (1983): "Scientific writing as social act: a review in the literature of the sociology of science." In P.V. Anderson, J. Brockmann, and C.R. Miller, eds., *New Essays in Technical and Scientific Communication: Research, Theory and Practice*. Farmingdale, New York: Baywood Publishing, 156-184.
- (1985): "Physicists reading Physics: schema-laden purposes and purpose-laden schema." *Written Communication* 2: 3-23.
- Berkentotter, Carol and Thomas Huckin (1995): *Genre Knowledge in Disciplinary Communication: Cognition/ Culture/ Power*. Hove: Lawrence Erlbaum Associates.
- Cooper, C. (1985): *Aspects of Article Introductions in IEEE Publications*. Unpublished MSc Dissertation, University of Aston in Birmingham.
- Francis, Gill. (1994): "Labelling discourse." In M. Coulthard, ed., *Advances in Written Text Analysis*. London and New York: Routledge, 83-101.
- Hewings, Ann. (1990): "Aspects of the language of Economics textbooks." In T. Dudley-Evans, and W. Henderson, eds., *The Language of Economics: The Analysis of Economics Discourse*. ELT Document 134, Modern English Publications in association with the British Council, 29-42.
- Hoey, Michael (1983): *On the Surface of Discourse*. London: Allen und Unwin.
- (1991): *Patterns of Lexis in Text*. Oxford: Oxford University Press.
- Huckin, T.N. (1987): *Surprise value in scientific discourse*. Paper presented at the CCC Convention, Atlanta.
- Hunston, Susan (1993): "Evaluation and ideology in scientific writing" In M. Ghadessy, ed., *Register Analysis: Theory and Practice*. London: Pinter, 57-74.
- (1994): "Evaluation and organisation in a sample of written academic discourse." In M. Coulthard, ed., *Advances in Written Text Analysis*. London and New York: Routledge, 191-218.
- Kaufer, David S. and Cheryl Geisler (1989): "Novelty in academic writing." *Written Communication* 6/3: 286-311.
- Knorr-Cetina, K.D. (1981): *The Manufacture of Knowledge*. Oxford: Pergamon.
- Latour, B. and Woolgar, S. (1979): *Laboratory Life: the Social Construction of Scientific Facts*. Beverly Hills: Cal. Sage Publications.
- McCarthy, Michael. (1990): *Vocabulary*. Oxford: Oxford University Press, 1990.
- Myers, Gregg (1990): *Writing Biology. Texts in the Social Construction of Scientific Knowledge*. Madison, Wisconsin: The University of Wisconsin Press.
- (1991): "Lexical cohesion and specialized knowledge in science and popular science texts." *Discourse Processes* 14: 1-26.
- Posteguillo, Santiago (1999): "The schematic structure of computer science research articles." *English for Specific Purposes* 18/2: 139-160.
- Swales, John (1990): *Genre Analysis: English in Academic and Research Settings*. Cambridge: England, Cambridge University Press.



- Thetela, Puleng (1997): "Evaluated entities and parametres of value in academic research articles." *English for Specific Purposes* 16/2: 101-118.
- Thompson, Geoff and Y.Y. Ye (1991): "Evaluation in the reporting verbs used in academic papers." *Applied Linguistics* 12: 365- 82.
- Winter, Eugene (1977): "A clause relational approach to English texts: a study of some predictive items in written discourse." *Instructional Science* 6/1: 1-92.
- Yearley, S. (1981): "Textual persuasion: the role of social accounting in the construction of scientific arguments." *Philosophy of the Social Sciences* 11: 409-435.

## Appendix

- Androutsos, D., K.N. Plataniotis and A.N. Venetsanopoulos (1999): "A novel vector-based approach to color image retrieval using a vector angular-based distance measure." *Computer Vision and Image Understanding* 75/1-2: 46-58.
- Assa, J., D. Cohen-Or and T. Milo (1999): "RMAP: a system for visualizing data in multidimensional relevance space." *The Visual Computer* 15: 217-234.
- Avrithis, Y., A.D. Doulamis, N.D. Doulamis and S. Kollias (1999): "A stochastic framework for optimal key frame extraction from MPEG video databases." *Computer Vision and Image Understanding* 75/1-2: 3-24.
- Barkan, E. and D. Gorden (1999): "The scan line principle: efficient conversion of display algorithms into scanline mode." *The Visual Computer* 15/2: 249-258.
- Campos, J., D. Donatelli and M. Silva (1999): "Structured solution of asynchronously communicating stochastic modules." *IEEE Transactions on Software Engineering* 25/2: 147-165.
- Chakraborty, S., K.Y. Yun and D.L. Dull (1999): "Timing analysis of asynchronous systems using time separation of events." *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems* 18/8: 1061-1077.
- Chang, S.C. and D.I. Cheng (1999): "Efficient Boolean division and substitution using redundancy addition and removing." *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems* 18/8: 1096-1106.
- Cukier, M., D. Powell, J. Arlat (1999): "Coverage estimation methods for stratified fault-injection." *IEEE Transactions on Computers* 48/7: 707-723.
- Emmel, P. and R. Hersch (1999): "Towards a color prediction model for printed patches." *IEEE Computer Graphics and Applications* July/ August: 54-60.
- Feo, J., S. Kahan and Z. Wu (1998): "Crash analysis on the Tera MTA." *IEEE Computational Science and Engineering* 5: 53-59.
- Fetzer, C. and F. Cristian (1999): "Fail-aware datagram service." *IEE Proceedings on Software* 146/2: 58-72.
- Fox, A.M., J.E. Cooling and N.S. Cooling (1999): "Integrated design approach for real time embedded systems." *IEE Proceedings on Software* 146/2: 75-85.
- Go, K. and N. Shiratori (1999): "A decomposition of a formal specification: an improved constraint-oriented method." *IEEE Transactions on Software Engineering* 25/2: 258-273.
- Goudreau, M.W., K. Kand, S.B. Rao, T. Suel and T. Tsantilas (1999): "Portable and efficient parallel computing using the BSP model." *IEEE Transactions on Computers* 48/7: 670-688.

- Holzleitner, L. and K.G. Mahmodu (1999): "Structural shape optimization using MSC/NASTRAN and sequential quadratic programming." *Computers and Structures* 70: 487-514.
- Jones, C.B., J.M. Ware and C.D. Eynon (1999): "Triangulated spatial models and neighbourhood search: an experimental comparison with quadtrees." *The Visual Computer* 15: 235-248.
- Kalitzin, S., B. Romeny and M. Viergever (1999): "Invertible apertured orientation filters in image analysis." *International Journal of Computer Vision* 31/2-3: 145-158.
- Knudsen, P.V. and J. Madsen (1999): "Integrating communication protocol selection with hardware/software codesign." *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems* 18/8: 1077-1095.
- Koenderink, J., J. van Doorn, K.J. Dana and S. Nayar (1999): "Bidirectional reflection distribution function of thoroughly pitted surfaces." *International Journal of Computer Vision* 31/2-3: 129-14.
- Manduchi, G. (1999): "Modelling a distributed real-time system for feedback control in a nuclear fusion experiment." *IEE Proceedings on Software* 146/2: 96-103.
- Martin, I.M. and D.C. Marinescu (1998): "Concurrent computation and daa visualization for spherical virus structure determination." *IEEE Computational Science and Engineering* 5: 40-52.
- Pinkston, T.M. (1999): "Flexible and efficient routing based on progressive deadlock recovery." *IEEE Transactions on Computers* 48/7: 649-669.
- Rashed, Y.F., M.H. Aliabadi and C.A. Brebbia (1999): "A boundary element formulation for a Reissner plate on a Pasternak foundation." *Computers and Structures* 70: 515-532.
- Sclaroff, S., M. La Cascia and S. Sethi (1999): "Unifying textual and visual cues for content-based image retrieval on the World Wide Web." *Computer Vision and Image Understanding* 75/1-2: 86-98.
- Segall, M., D. Reinstein and N. Johnson (1999): "Visualizing VHF ultrasound of the human cornea." *IEEE Computer Graphics and Applications* July/ August: 74-82.
- Song, X. and J.D. Baldwin (1999): "A novel node-based structural shape optimization algorithm." *Computers and Structures* 70: 569-581.
- Sun, Y., F. D. Fracchia, T.W. Calvert and M.S. Drew (1999): "Deriving spectra from colors and rendering light interference." *IEEE Computer Graphics and Applications* July/ August: 61-66.
- Varvarigou, T.A., M.E. Anagnostou and S.R. Ahua (1999): "Reconfiguration models and algorithms for stateful interactive processes." *IEEE Transactions on Software Engineering* 25: 401-415.
- Weickert, J. (1999): "Coherence-enhancing diffusion filtering." *International Journal of Computer Vision* 31/2-3: 111-127.
- Zhu, S. and P.J. McKerrow (1998): "Visualizing reflection of a curved surface." *IEEE Computational Science and Engineering* 5: 30-38.