UNIVERSIDADE DE SÃO PAULO

WRITING TOOLS AND A SOFTWARE ARCHITECTURE TO ASSIST WRITING IN A FOREIGN LANGUAGE

SANDRA M. ALUÍSIO
MARIA C.F. DE OLIVEIRA.
Á LVARO GARCIA NETO
PAULO C. MASIERO
OSVALDO N. OLIVEIRA JR.

No. 27

NOTAS

Instituto de Ciências Matemáticas de São Carlos

Instituto de Ciências Matemáticas de São Carlos

ISSN - 0103-2577

WRITING TOOLS AND A SOFTWARE ARCHITECTURE TO ASSIST WRITING IN A FOREIGN LANGUAGE

SANDRA M. ALUÍSIO
MARIA C.F. DE OLIVEIRA.
Á LVARO GARCIA NETO
PAULO C. MASIERO
OSVALDO N. OLIVEIRA JR.

No. 27

NOTAS DO ICMSC
Série Computação

São Carlos
Mai./1996

Writing Tools and a Software Architecture to Assist Writing in a Foreign Language

Sandra M. Aluísio[1], Maria C.F. de Oliveira[1], Álvaro Garcia Neto[2],

Paulo C. Masiero[1], Osvaldo N. Oliveira Jr.[2]

[1]Instituto de Ciências Matemáticas de São Carlos
Universidade de São Paulo
CP 668 - 13560-970 - São Carlos - SP - BRAZIL

[2]Instituto de Física e Química de São Carlos
Universidade de São Paulo
CP 369 - 13560-970 - São Carlos - SP - BRAZIL
FAX: +55 162713616
e-mail: chu@ifqsc.sc.usp.br

Abstract

This article describes the AMADEUS writing environment, which is aimed at providing tools to assist the production of scientific texts in English. The intended audience for AMADEUS are non-native users of English who must write scientific documents. A generic model is proposed for developing writing tools for these target users, which is based on the reuse of linguistic material within a case-based reasoning approach that also includes interactive adaptations. The power of the method has been demonstrated in developing a Support Tool specific for Introductory Sections of Experimental Physics, whose architecture is sufficiently generic to allow extension to other sections of a paper and other domains. We also discuss the integration of AMADEUS into a conceptual distributed software architecture that enables its tools to intercommunicate through a network.

1. Introduction

The use of computers in writing goes back to ancient times in Computer Science, as text editors and text formatters were already available in the very early computer environments. Word processing programs grew out to become perhaps the most popular application of personal computers, reaching widespread use. These programs were later enhanced by spelling checkers, and used in conjunction with idea processors (outliners and structure editors), thesauruses, style and grammatical analysers, and a number of other writing aids. Despite the wide availability and usefulness of these tools, they fail to address "the large scale structure of the writing task" [Sharples-92, Genthial-92, Williams-89]. New tools are expected to contemplate the integration of already existing facilities, plus other innovative tools, into Computer Aided Writing (CAW) environments which are prepared to assist the writer throughout the writing process. This is an essential requisite if writing support technology is to evolve in such a way as to meet increasing users' demands and expectations.

This work describes a proposal for a CAW environment (AMADEUS - **AM**iable **A**rticle **D**evelopment for **U**ser **S**upport) targeted at assisting non-native users of English (nnuE) within academic environments. Writing tools are presented that provide direct support in solving specific difficulties faced by (Brazilian) writers when producing texts in English. Also proposed is a distributed software architecture for integrating a number of modules addressing different aspects of the writing task. This conceptual architecture is presented in Section 4, following a brief description of some noteworthy CAW environments in Section 2, and a report on the development of a linguistic aid module for AMADEUS in Section 3.

2. Computer Aided Writing (CAW) Environments

A number of writing tools have been described in the literature whose ultimate goal is to improve the quality of the texts produced by different groups of users on different settings. In this section we try to summarize current research in the CAW field, though this brief review is by no means comprehensive. Writing tools are usually classified into **Pre-processing tools**, **Post-processing tools** and **Environments**, which encompass both pre- and post-processing tools. The post-processing tools are the most widely investigated, and several of them are commercially available, such as Grammatik [RSI-92] and Correct Grammar [WTG-91]; other examples are listed in Table I. These tools generally offer statistic- and rule-based information for correcting spelling, style, text readability and grammar.

The pre-processing tools, also called outliners, idea processors or organizers, are aimed at helping writers to select and organise the material to be written. Two examples, *viz.* Boxweb and Story Space, are mentioned in Table I. Many of the currently available text processors provide an outliner. This simple but powerful facility helps writers to identify hierarchical relationships in their raw material and allows them to expand or to hide the different organisation levels of the text, thus reducing the cognitive burden of the writing process.

The main goal of the environments is to support the writer throughout the different phases of the writing process. They consist of pre- and post-processing tools which are preferably integrated within a cognitive writing

model. Table I mentions several such environments, from which WE, The Writer's Assistant and CSile are representative examples. **WE** (Writing Environment) is a CAW environment based on careful consideration of the cognitive processes employed by writers during text production [Smith-88]. Four main phases of the writing process were identified, namely exploratory, organizational, writing and editing. The system provides four operating modes, each targeted at one of these phases.

The **Writer's Assistant** [Sharples-90] is another system whose aim is to provide an integrated CAW environment for the preparation (planning and design) of complex documents, such as research reports and technical documentation. The objective of providing cognitive support for writers is accomplished by giving them three possible views of the document: a linear view, a structure view and a notes network view. These views can be displayed in different presentation formats, to suit users' needs and focus of attention.

Genthial and Courtin [Genthial-92] report on the development of a CAW environment for French writers whose main feature is a distributed architecture comprised by independent modules which share a minimum set of data structures. The system modules include a text processor, a multipurpose lexicon, a syntactic checker, a structure manager, a morphological module and others.

3. Tools for scientific writing by non-native users of English

Due to the diversity of writing tasks, writers abilities and document styles, most CAW systems concentrate on specific classes of writers and documents. A general-purpose software able to present generalized solutions to problems facing any kind of writer is not a realistic assumption [Williams-89], at least in the current state of knowledge. A restricted-class problem of widespread interest is the production of scientific articles. This problem has two distinctive and interesting characteristics: **Scientific articles are highly structured**, being thus amenable to treatment by current technology; and **scientific papers are frequently written by non-native users of English (nnuE)**, since English is the *lingua franca* of the scientific community. The environment we propose is targeted at writers producing scientific texts in academic environments. Software tools to assist in the learning of languages by foreigners [Byron-90, Lelouche-92] do not address specifically the production of quality scientific texts.

3.1 Generic Principles for Developing Tools to Assist Writing in a Foreign Language

The process of writing has been recognized as a cyclic process in which generation and organization of ideas (planning) are alternated with the linguistic realization and editing/revising stages [Hayes-80, Bruce-78]. We propose an approach for developing writing tools to assist nnuE in preparing the first draft of a scientific paper which can cover the different stages of writing. This approach is based on four main principles which are presented below. As we shall see, in actual implementations of less sophisticated tools only 2 or 3 of these principles need to be applied. The principles are:

1) Reuse of non-factual linguistic material of authentic texts as a way of minimising mother tongue interference.
2) Pre-compilation of knowledge of a particular genre into a detailed schematic structure which is used for mapping sections of a paper into cases. This structure also serves as a knowledge source for adapting cases according to the user needs.
3) Use of Case-Based Reasoning for modeling the various stages of the writing process, including interactive adaptations.
4) Use of Critiquing systems with due consideration of distinct needs of different classes of users.

The rationale behind these principles is discussed in the following. Several difficulties faced by nnuE in scientific writing are related to the absence (or misuse) of collocations and expressions normally employed by competent native writers [James-84]. This is usually caused by the so-called mother tongue interference. A possible way to overcome such a difficulty is to **re-use non-factual linguistic material** from authentic texts belonging to the same genre of the text under consideration. Reusable texts have in fact been exploited for producing standardized documents [Born-92], for reducing the duration of writing tasks [Buchanan-92] and in natural language generation or interpretation systems using phrasal lexicons [Kukich-83, Jacobs-85, Hovy-90, Smadja-91]. A phrasal lexicon for any particular application requires extensive corpus analysis for the **pre-compilation of the knowledge of the genre into a detailed schematic structure**. A writing tool has been developed using these two principles [Fontana-93] which provided encouraging results for some non-native writers. However, as discussed in the next sub-section, the tool failed to help less experienced writers with poorer knowledge of English. These users had difficulties in identifying expressions appropriate to their needs and also in making sensible choices of cohesive links for building up a paragraph.

To help overcome the problems of cohesive links, and to a lesser extent, problems in coherence, the philosophy of **case-based reasoning together with interactive adaptations** [Kolodner-92] was suggested as the third principle for developing writing tools. The phrasal lexicon is to be stored in a case base, where each case corresponds to a coherent piece of text. A case not only illustrates the rhetorical structure of the text but also instantiates the use of cohesive devices using real text. Since the rhetorical structure is apparent, cases can be retrieved using a pre-defined metrics based on structural similarity with the user requisition. **Interactive adaptations** can then be performed using one (or more) of the cases returned to the user. Four revision operations are envisaged: 1) changes in the lexical and/or syntactic material of a rhetorical message[1]; 2) changes applied to the selected rhetorical strategies recovering similar ones; 3) addition of messages to a specific strategy; 4) deletion of messages, the opposite operation to 3). Operation 1) consists simply of replacement of linguistic material within a template, while operation 2) may affect the rhetorical information conveyed by the strategy. The revision operations allow the user to make use of linguistic material from other cases from the base, which turns the system extremely

[1]Rhetorical messages refer to the different types of information conveyed in a rhetorical strategy.

flexible. The use of this third principle is instantiated in the next sub-section where a Support Tool for writing Introductions of Experimental Physics papers is presented.

Tools developed using the principles outlined above, sophisticated as they might be, have no control over the quality of the draft they helped the user to produce. There is no mechanism to check whether the choices made by the user were reasonable or even check consistency among the rhetorical sub-strategies comprising a rhetorical strategy. While some consistency checks may be easily implemented, it is possible that the tools may need to include knowledge of why writers employ rhetorical strategies in specific conditions which obviously requires further research into the writing process. The **critiquing systems** available in the literature (e.g. [Silverman-92, Rankin-93]) may serve as a starting point for the development of such tools. One of the versions proposed for the AMADEUS linguistic module, called **Tutorial**, is to be developed employing the fourth in addition to the other 3 principles.

With regard to the user-system interaction, within the proposed approach it is to be carried out in three stages:

i) The user provides information on the rhetorical structure of the intended introduction;
ii) The system returns a set of best-match cases, for selection by the user;
iii) The user adapts the selected case(s) to produce his/her own text.

3.2 The AMADEUS linguistic aid Module

This module helps users to generate a first draft of a scientific paper. Because distinct users may have different needs, more than one version were conceived for this writing tool. The first module, which is already fully implemented for a subfield of Experimental Physics, is called the **Reference Tool**. Historically, this was developed before the four generic principles described above were established, and in fact analysis of the results from applying the Reference Tool to Brazilian post-graduate students writing in English were essential in developing the rationale for the generic approach. It employs principles 1) and 2) of Section 3.1 through an empirical resource described in [Oliveira-91] which consists in providing writers with input material in the form of useful contextualized expressions compiled from scientific papers and books. Template expressions are stored in a sentences base with blank spaces (representing missing words) where reference to specific material is expected. Users may access templates in the sentences base in one of three ways:

1. Taxonomy. The expressions are indexed according to a taxonomy which is based on the main topics likely to appear in a paper.

2. Communicative Goals. The search identifies those expressions which are related to a specific communicative goal (describe, justify, compare, contrast, etc.) [McKeown-85].

3. *Keywords*. The search is carried out by selecting a keyword that appears frequently in texts of a given area.

The Reference Tool also possesses an Acquisition Mode which enables customisation of the sentences database by the inclusion of new expressions and/or new items in the taxonomy. The use of this tool has been reported in [Fontana-93]. In addition to offering authentic input for academic writing at different degrees of complexity, the tool helped students to get started in their writing tasks. This is encouraging because getting over the writer's block has been mentioned as one of the most important problems faced by writers [Williams-90]. The most critical shortcoming of the tool, on the other hand, has been the frustration by users who had too high expectations and believed the tool could solve most of their writing problems. Students with poor knowledge of English had difficulties in using the resource because they were not able to make sensible connections between different expressions, and also because they normally made far too many mistakes when adding in their own text.

For minimising these problems, the philosophy of case-based reasoning together with interactive adaptations was adopted on the second type of tool - the **Support Tool**. The power of the approach was demonstrated by implementing a prototype for Introductory Sections of Experimental Physics papers. The case base contains 54 instances of authentic Introductory sections from two well-regarded periodicals: the Physical Review Letters and Thin Solid Films. The corpus analysis of these 54 introductions enabled the establishment of a schematic structure which may be considered as typical of the research area [Aluisio-95a]. As far as the main components are concerned, the analysis confirmed what has been proposed by Weissberg [Weissberg-90] and Swales [Swales-90] in that the most frequent components of Introductions are *setting, review, gap, purpose, methodology, results, value and layout*. In addition, 30 rhetorical strategies were identified which are used to organise 45 different types of information (called here as messages). Each strategy is realized linguistically by 2 or 3 messages. Fig. 1 shows the main components and strategies identified.

Each case (whole Introduction) is represented in a computational formalism with the corresponding text associated. A case named TSF4 is illustrated in Fig. 2 whose list of components includes: setting, review, gap and two types of purpose. The fragments of underlined text correspond to the reusable parts. The sentences are numbered since they are the basic unit of analysis. The factual linguistic material appears as normal text. The representation of the schematic structure was based on [McKeown-85] and [Hovy-92]. It is through this network that the four rules of interactive adaptations can be employed in the last stage of the writing process (the revision/editing process). The corresponding Prolog representation of TSF4 is shown in Fig. 3.

The prototype was implemented in Sicstus Prolog and its interface developed using the XView Toolkit [Heller-90] and the Netscape browser (Fig. 4). In the Gathering of Features stage the user is offered several menus from which he/she should select the intended components and strategies. Based on the user choices, the system applies the similarity metrics and four best-match cases are returned. Three ways of pattern matching are used:

perfect match (equal lists), proper undermatch (sublist) and non-proper undermatch (intersection) (see [Aluisio-95b]). The user can then work on one (or more) authentic introductions by transferring non-factual linguistic material into his/her own working area on the right hand side of the window. Assistance is also provided in the editing process through the four revision operations mentioned in the previous section, as shown on the top part of the screendump of Fig. 4.

4. The AMADEUS software architecture philosophy

Even with the restriction of scope, the task of helping writers is still awesome. Several of the tools still require advances in the technology of Natural Language Processing, Artificial Intelligence and others. It is unwise, therefore, to create monolithic tools to deal with this problem. A better approach is to define a software architecture to provide a supporting harness in which a set of separate independent tools can be integrated. AMADEUS is being conceived as a distributed system with a collection of modules ranging from pre-processing to post-processing tools. It should therefore feature as an environment according to the classification of Table I. To enable the exchange of information amongst the (in-house developed or off-the-shelf) modules, an interconnection protocol shall be provided. The interconnection protocol will also ensure that tools may be invoked by users from different platforms in a totally transparent manner. The philosophy underlying the AMADEUS environment has the ultimate objective of providing a software architecture with the following characteristics:

1. A well-defined intercommunication protocol, enabling any tool to transfer data and actions to any other tool within the architecture.
2. A flexible set of text processing tools, that may be invoked in any number or order from the available set, and include thesauruses, spelling checkers, false cognate detectors, template-based linguistic aids, reference and cross-reference generators, bibliography database management systems, jargon detectors and correctors, statistic index calculators, document structurers, etc.
3. A high degree of interoperability, due to the abilities of reading and writing different file formats, of communicating amongst tools and of accessing the tools through any convenient user interface.
4. A collection of graphical user interfaces displaying suitable metaphors for writers.

In addition to sophisticated word processing or desktop publishing functions, it is reasonable to expect also from a writing environment assistance in the selection and organisation of the material to be included in the text, as well as on how to express this material in sentences and paragraphs. We might then be heading towards the development of more specialised tools which address the needs of specific classes of writers. The interfaces of such specialised tools should be designed to consider not only the specific goals of the tool; e.g. whether it is aimed at supporting or teaching its users, but also their degree of expertise at the writing process and their proficiency in the

language in which the text is being written. In this context, it would be useful if the software could pickup common linguistic mistakes made by non-natives (e.g. the use of an incorrect preposition). Also, users should be able to search for the English synonymous of a given word.

We believe that a computer environment for assisting nnuE in the preparation of technical/scientific documents in English must include a number of independent but integrated tools targeted at the preparation, compilation and post-processing analysis of texts. The proposed AMADEUS environment relies on a distributed architecture providing a collection of tools compliant to a connection protocol enabling the exchange of information between them, as shown in Fig. 5. An important aspect of this architecture is the facility provided for the writer to "pick and choose" those tools considered to be more adequate for the task in hand. This is so because it is unreasonable to expect that a particular tool or environment can suit every user and task, as the nature of help required by a writer when producing a scientific paper varies enormously according, among other factors, to his/her expertise in English and experience in writing under scientific register, the level of cross-referencing required by the subject matter, the recurrence to specific jargon, etc.

Fig. 5 also shows that there are two types of application: those that comply with the intercommunication protocol and those that do not. It is expected that in-house developed applications will comply with the protocol, while commercial applications will obviously not. These will address the underlying layers through intervening software drivers. The interconnection protocol also ensures that users can access tools from different platforms transparently. The protocol must provide each tool on the environment with the capability of transferring data and communicating actions to all other tools, either directly or through intervening software drivers. In order to do this, the protocol draws heavily on established communication protocols and also on established interface definitions, such as the X graphics protocol.

4.1 Available and proposed tools

Possible tools to be integrated into the environment include:

Text editors and wordprocessors. We wish to let users select which software to use from several possible alternatives, as the choice of a wordprocessor or equivalent is greatly a matter of habit and personal taste. Besides, considerable effort has been put on the look-and-feel and usability of modern wordprocessors, and this investment should not be wasted.

Multipurpose lexicons. English to English and bilingual (in our case, Brazilian Portuguese to English, English to Brazilian Portuguese) dictionaries.

Thesaurus. Preparation of electronic cross-cultural thesauruses is a difficult linguistic task that must be, at some stage, undertaken.

Spelling checker. Spelling checkers have become widely available and must be contained in any writing environment.

Syntax checker. This tool should take into account the specific problems of writers with a certain mother tongue, in our case, Brazilian Portuguese.

Document style analyser. A large-grain text analyser specific for the production of scientific texts must be provided. This tool analyses the placement of sections taking into account the cognitive goals of each chunk of text. It attempts to flag inconsistencies such as the introduction of new concepts in the conclusion section.

False cognate filter. False cognate words plague most English texts produced by non fluent writers with a background on Romance tongues.

Template based linguistic aid. This tool corresponds to the software implementation described in the previous Section.

Bibliography and citation tools. Several tools are already available, but new generation tools should address issues such as groupware, internal document cross-referencing and integrated networking for bibliographic research into libraries.

These modules, providing different levels of linguistic aid, shall be presented to users by a custom user interface which provides an adequate metaphor for writers - we consider the desktop metaphor adopted in current graphical interfaces inadequate, at a higher level, to many writers - and be supported by a software architecture that fully integrates the different writing tools. In such an architecture, the user interface can be viewed as another application and should therefore be "transparent". That is, the writer would have the look-and-feel of whatever Operating System (OS) interface he/she is using. On the other hand, users may prefer an interface providing, for instance, a more convenient metaphor for writers, which would then "mask" the underlying OS interface.

5. Conclusions

Despite the wide availability of tools for supporting the writing task, there has been little concern in developing CAW environments addressing the particular problems faced by non-native writers. Obvious exceptions are the grammar checker ISCA [Bolt-95] and the recent initiative in the Composer Project [Pemberton-95] which is aimed at guiding and critiquing students' academic work. In this paper we propose generic principles for developing writing tools for these target users. Reuse of non-factual linguistic material is effected within a case-based reasoning approach where each case consists of a whole section of a paper (Introduction for the implemented Support Tool described here). Revision operations allow interactive adaptations which cater for the user needs and lend great flexibility to the writing tool. Because writing environments must take into account the writing process as a whole in addition to the users' specific difficulties, we propose a software architecture that can provide an

integrated environment for non-native writers. Further work should then concentrate on the development of more sophisticated tools and integration into the environment, which certainly requires the establishment of a multidisciplinary team.

Acknowledgments

The authors acknowledge the financial assistance from CNPq and FAPESP.

References

[Aluisio-95a] Aluisio, S.M.; *Ferramentas para Auxiliar a Escrita de Artigos Científicos em Inglês como Língua Estrangeira*, PhD Thesis, University of São Paulo (Brazil), 1995.

[Aluisio-95b] Aluisio, S.M. and Oliveira Jr., O.N.; A Case-Based Approach for Developing Writing Tools Aimed at Non-native English Users, Lecture Notes in Artificial Intelligence, to appear.

[Bolt-95] Bolt, P.; A Grammar Checking Program: ISCA, *Proc. of VIII Conference on Writing and Computers*, (Eds. Preston, C. and Scott, M.), September, London, 1995.

[Born-92] Born, G.; A Hypertext-Based Support Aid for Writing Software Documentation, *Computers and Writing - State of the Art*, (Eds. Holt, P.O. and Williams, N.), pp. 266-277, Intellect, Oxford, 1992.

[Bruce-78] Bruce, B., Collins, A., Rubin, A., and Gentner, D.A.; *Cognitive Science Approach to Writing*, Center for the Study of Reading, Technical Report No. 89, 1978.

[Buchanan-92] Buchanan, R.A.; Textbase Technology: Writing with Reusable Text, *Computers and Writing - State of the Art*, (Eds. Holt, P.O. and Williams, N.), pp. 254-265, Intellect, Oxford, 1992.

[Byron-90] Byron, G.; eL: A Tool for Language Learning. *CALL*, Vol. 2, pp. 83-91, 1990.

[Fontana-93] Fontana, N.M., Caldeira, S.M.A., De Oliveira, M.C.F. and Oliveira Jr., O.N.; Computer Assisted Writing - Applications to English as a Foreign Language, *CALL*, Vol. 6, pp. 145-161, 1993.

[Genthial-92] Genthial, D. and Courtin, J.; From Detection/ Correction to Computer Aided Writing. *Proceedings of the Coling-92*, Nantes, pp. 1013-1018, 1992.

[Hayes-80] Hayes, J.R. and Flowers, L.S.; Identifying the Organization of Writing Processes, *Cognitive Processes in Writing*, (Eds. Gregg, L.W. and Steinberg, E.R.), pp. 3-30, Hillsdale, N.J., 1980.

[Heller-90] Heller, D. *XView Programming Manual.* O' Reilly & Associates, Inc., 1990.

[Hovy-90] Hovy, E.; Pragmatics and Natural Language Generation, *Artificial Intelligence*, Vol. 43, pp. 153-197, 1990.

[Hovy-92] Hovy, E., Lavid, J., Maier, E., Mittal, V. and Paris, C.; Employing Knowledge Resources in a New Text Planner Architecture, *Lecture Notes in Artificial Intelligence*, Vol. 587 (Eds. Dale, R., Hovy, E., Rosner, D. and Stock, O.), Springer-Verlag, 1992.

[Jacobs-85] Jacobs, P.S., PHRED: A Generator for Natural Language Interface, *Computational Linguistics*, Vol. 11, pp. 219-242, 1985.

[James-84] James, K.; The Writing of Theses by Speakers of English as a Foreign Language: The Results of a Case Study, *ELT Documents: 117, in Common Ground: Shared Interests in ESP and Communication Studies*, Williams et alii (eds), Pergamon Press, 1984.

[Kolodner-92] Kolodner, J.L.; An Introduction to Case-Based Reasoning, *Artificial Intelligence Review*, 6, pp. 3-34, 1992.

[Kukich-83] Kukich, K.; *Knowledge-based Report Generation: A Knowledge Engineering Approach to Natural Language Report Generation*, PhD Thesis, University of Pittsburg, 1983.

[Lelouche-92] Lelouche, R.; Using Multiple Knowledge Bases to Help Teach Some Pragmatic Aspects of French. *CALL*, Vol. 4, pp. 29-40, 1992.

[McKeown-85] McKeown, K.R.; Discourse Strategies for Generation Natural-Language Text. *Artificial Intelligence*, 27, pp. 1-41, 1985.

[Oliveira-91] Oliveira Jr., O.N, Caldeira, S.M.A. and Fontana, N.M.; Chusaurus: A Writing Tool Resource for Non-Native Users of English. *Proceedings of the XI International Conference of the Chilean Computer Science Society*, pp. 59-70, 1991.

[Pemberton-95] Pemberton, L., Shurville, S., Broady, E., Drake, S., Hartley, T. and Salkie, R.; Development Methodology for a Computer Based Support Tool for Academic Writing in a Second Language. *Proc. of VIII Conference on Writing and Computers*, (Eds. Preston, C. and Scott, M.), September, London, 1995.

[Rankin-93] Rankin, I.; Natural Language Generation in Critiquing. *The Knowledge Engineering Review*, Vol. 8, pp. 329-347, 1993.

[RSI-92] Reference Software International, *Grammatik 5 User's Guide*. Reference Software International, Washington DC, USA, 1992.

[Sharples-90] Sharples, M. and Pemberton, L.; Guidelines for the Design of User-centred Document Processors. *CALL*, Vol. 2, pp. 37-57, 1990.

[Sharples-92] Sharples, M. and Pemberton, L.; Representing Writing: External Representations and the Writing Process, In *Computers and Writing - State of the Art*, (Eds. Holt, P. and Williams, N.), pp. 319-336, Intellect, Oxford, England, 1992.

[Silverman-92] Silverman, B.C.; Survey of Expert Critiquing Systems: Practical and Theoretical Frontiers, *Communications of the ACM*, 35, pp 107-127, 1992.

[Smadja-91] Smadja, F.; Retrieving Collocational Knowledge from Textual Corpora. An Application: Language Generation, PhD Thesis, University of Columbia (USA), 1991.

[Smith-88] Smith, J.B. and Lansman, M.; *A Cognitive Basis for a Computer Writing Environment*. Technical Report. Chapel Hill, NC:University of North Carolina, Department of Computer Science, 1988.

[Swales-90] Swales, J.; *Genre Analysis - English in Academic and Research Settings*, Cambridge University Press, 1990.

[Weissberg-90] Weissberg, R. and Buker, S.; *Writing Up Research: Experimental Research Report Writing for Student of English*. Prentice Hall Regents, Englewood Cliffs, NJ, 1990.

[Williams-89] Williams, N.; Approaches to Computer Assisted Writing. In *Computers and Writing: Models and Tools*, (Eds. Williams, N. and Holt, P.), pp. 1-15, Intellect, Oxford, England, 1989.

[Williams-90] Williams, N.; Writers' Problems and Computer Solutions. *CALL*, Vol. 2, pp. 5-25, 1990.

[WTG-91] Writing Tools Group, Inc.; *Using Correct Grammar for Windows*. Writing Tools Group, Inc., Sausalito, California, USA, 1991.

Table I - Writing Tools and Environments

Tool	Developers Group	Classification	Features
Boxweb	P. Wayner Cornell University	Pre-processing	High-level language for creating outlines
Story Space	Eastgate Systems Inc., Cambridge, USA	Pre-processing	Graphical ideas organizer, outliner, hypertext browser
APGS (Audit Program Generator System)	R. Buchanan Boeing Computer Services	Environment	Production of technical documents by reusing texts
SESS (Software Engineering Support System)	G. Born SD Scicon	Environment	Production of standard documents with knowledge-based support
CONST-Project	J. Beeken University of Leuven	Environment	production and analysis of commercial and business texts
Writer's Assistant	M. Sharples University of Sussex	Environment	Computer-based cognitive support system for creating complex documents
WE (Writing Environment for Professionals)	J.B. Smith University of North Carolina at Chapel Hill	Environment	Highly visual system based on a cognitive model for written communication
CSILE (Computer- Supported Intentional Learning Environments)	M. Scardamalia Ontario Institute for Studies in Education, Canada	Environment	A communal database designed to encourage group processes and progressive discourse
SEPIA (Structured Elicitation and Processing for Authoring)	N. Streitz GMD-IPSI, FRG	Environment	Active, knowledge-based authoring tool for creating argumentative texts
Genthial's CAW Environment	D. Genthial Laboratoire de Génie Informatique, France	Environment	Coherent set of tools aimed at production, maintenance, edition and communication of texts
Ruskin	CIRG (Comm. and Inform. Research Group) Sheffield City Polytechnic	Post-processing	Analyser of texts at many different linguistic levels
EPISTLE	G. Heidorn IBM	Post-processing	Grammar and style checking of business communication
McRuskin	S. McGovern Heriot-Watt University	Post-Processing	Enhanced version of Ruskin with discourse analysis
Writer's Workbench	N. MacDonald Bell Lab's	Post-processing	Collection of separated tools for correcting spelling, grammar and style

C1: Setting
- S1 Introducing topics of a given area
- S2 Familiarizing with terms, objects or processes
- S3 Argumenting about proeminence of topic/area

C2: Review
- S1 Historical review of the area
- S2 Current trends
- S3 Citations grouped from general to specific topic
- S4 Progress in the area
- S5 Requirements for the progress
- S6 State of the art
- S7 Citations and cyclic gaps
- S8 Citations grouped according to approaches

C3: Types of Gap
- S1 Existing conflicts or unresolved problems
- S2 Limitations of previous works
- S3 Issues not considered yet

C4: Purpose
- S1 Indicating the main purpose
  - S1A Resolving a conflict
  - S1B Presenting a new approach, or methodology or technique
  - S1C Presenting improvements within a particular topic
  - S1D Extending the author´s own previous work
  - S1E Proposing an alternative approach
  - S1F Presenting a comparative work
- S2 Specifying the purpose
- S3 Introducing further purposes

C5: Methodology
- S1 Indicating criteria and conditions
- S2 Describing materials and methods
- S3 Justifying choices of methods and materials

C6: Main Results
- S1 Presenting results
- S2 Commenting on the results

C7: Value of the Research

C8: Structure of the article
- S1A Indicating sections
- S1B Listing issues to be addressed

Figure 1. Main Components, C, and Strategies, S, comprising the detailed schematic structure of an Introduction

SETTING: FAMILIARIZING TERMS OR OBJECTS OR PROCESSES
1) Conductive LB films of charge transfer complexes based on TCNQ, TTF and related donors [1], are characterized by strong one-dimensional interactions and are therefore susceptible to stacking defects and the Peierls instability.
REVIEW: REQUIREMENTS FOR THE PROGRESS IN THE AREA
2) It should be possible to overcome such defects and thus produce films of higher conductivity if a complex of higher dimensionality is chosen.
3) Indeed, Nakamura et al. [2] have shown that monovalent salts of a range of long-chain alkylammonium salts of M(dmit)2, where M = Au, Ni, Pd, or Pt may be deposited as LB films and oxidised with bromine or iodine to yield molecular films with conductivities in the range 0.001-30S/cm.
GAP: UNRESOLVED CONFLICT OR PROBLEM AMONG PREVIOUS STUDIES
4) However, owing to their instability at the air-water interface, it was necessary in most cases to deposit the complexes as 1:1 mixtures with icosanoic acid.
PURPOSE: PRESENTING AN EXTENSION OF A PREVIOUS AUTHOR'S WORK + INTRODUCING MORE PURPOSES
5) § In a preliminary report [3], we have shown that it is possible to deposit LB films of the pure didodecyldimethylammonium salts of the divalent metal complexes [M(dmit)2]2- and [M(mnt)2]2-.
6) In this report we describe further investigations into the behavior of these compounds (complexes 1 and 2, respectively) at the air-water interface and extend our studies to the monovalent salt (complex 3) for comparison.
(Taylor,D.M.et alli. *Monolayer characterization and multilayer deposition of conducting Langmuir-Blodgett films*, Thin Solid Films, 210/211 (1992) pp.287-289)

Fig. 2. One of the Introductions of the case base

```
case( tsf4,
    [c(setting,s(familiarizing_terms_or_objects_or_processes,
                    [m(semi_formal_definition, tsf4,1)])),
     c(review,s(requirements_for_the_progress_in_the_area,
                    [m(topic_complexity, tsf4,1),
                     m(requirements_progress,tsf4,2),
                     m(evidence,tsf4,3)])),
     c(gap,s(unresolved_conflict_or_problem_among_previous_studies,
                [m(conflict_or_problem,tsf4,4), m(cause,tsf4,4)])),
     c(purpose,s(presenting_an_extension_of_a_previous_authors_work,
                [m(author_prominent,tsf4,5), m(extension,tsf4,6)])),
     c(purpose,s(introducing_more_purposes, [m(addition,tsf4,6)]))], _ ).
```

Fig. 3. Rhetorical features represented in Prolog

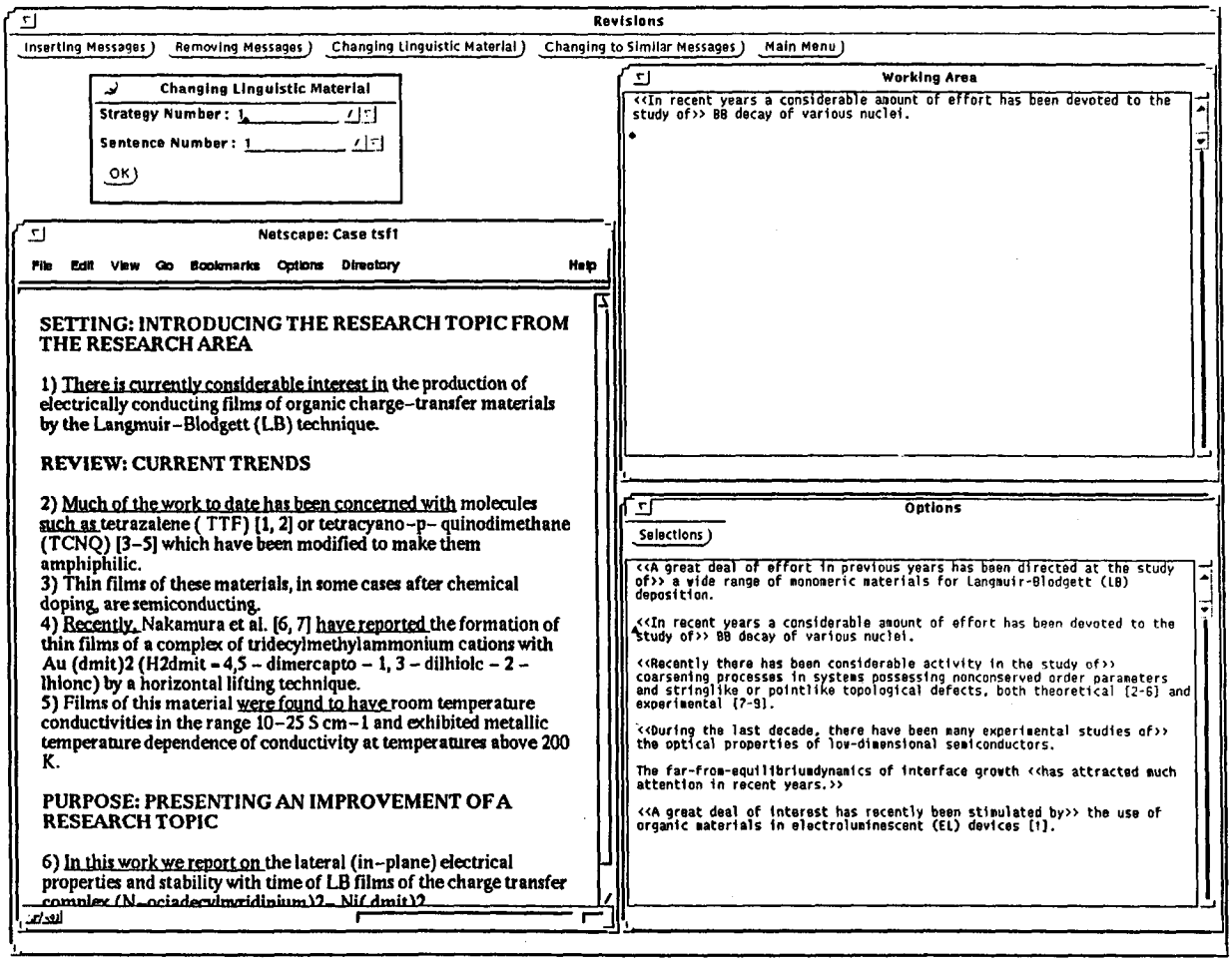

Fig. 4 Screendump of the Support Tool interface at the revision/editing stage. A selected case is shown on the left-hand side window which is to be modified using the four revision operations represented on the top of the screen. The options contained in the lower window were obtained from other cases of the base. Any material selected can be inserted into the working area by the clicking of a button.

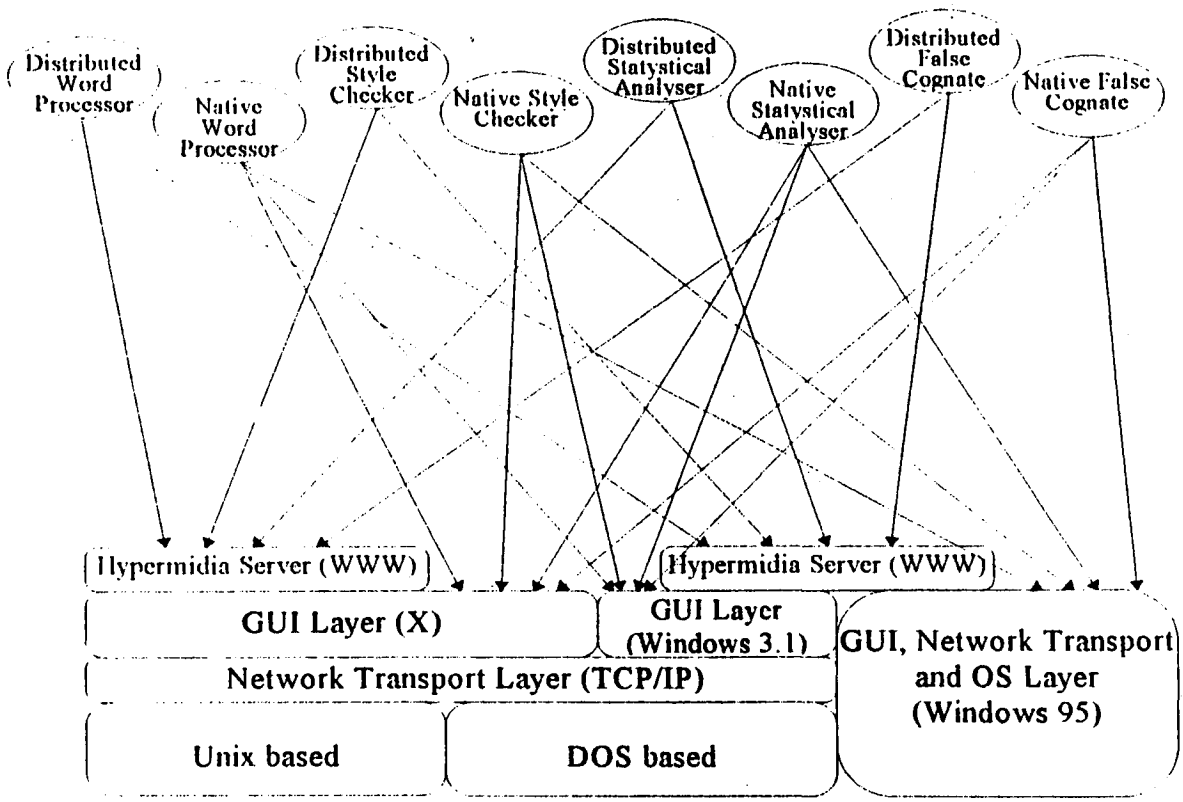

Figure 5 - A possible architecture for the Amadeus Environment

NOTAS DO ICMSC

SÉRIE COMPUTAÇÃO

026/96 ALUISIO, SANDRA M.; OLIVEIRA JR., OSVALDO N. - A detailed schematic structure of research paper introductions: an application in support-writing tools.

025/96 NUNES, M.G.V.; HASEGAWA, R.; KAWAMOTO, S.; OLIVEIRA, M.C.F. DE; TURINE, M.A.S.; GHIRALDELO, C.M.; OLIVEIRA JR., O.N.; RIOLFI, C.R.; SIKANSKI, N.S.; MARTINS, T.B. - Style and grammar checkers for brazilian portuguese.

024/96 FORTES, R.P.M. - Uma ferramenta orientada a links para avaliação de hiperdocumentos.

023/96 CANSIAN, A.M.; MOREIRA, E.S.; MAURO, R.B.; MORISHITA, F.T.; CARVALHO, A.C.P.L.F. - Um sistema adaptativo de detecção de intrusão em redes de computadores.

022/96 BRIGANTE, W.J.; MOREIRA, E.S. - Utilização de Monitores OLTP no gerenciamento de ambientes de manufatura integrada voltados à produção discreta.

021/95 BEZERRA, L.A.F.; SANTANA, R.H.C.; SANTANA, M.J. - Sistema auxiliar de arquivos baseado em disco WORM para ambientar computacional distribuído.

020/95 NUNES, M.G.V.; HASEGAWA, R. - PROTEMA: intelligent tutoring systems for mathematics.

019/95 OLIVEIRA, M.C.; TURINE, M.A.S.; MASIERO, P.C. - A statechart - based model for hypertext.

018/95 PIMENTEL, M.G.C. - Alternative operations for browsing hypertext.

017/94 ROMEIRO, N.M.L.; CASTELO FILHO, A. - Análise Comparativa de Métodos Numéricos de equações algebrico-diferenciais.