

This item was submitted to [Loughborough's Research Repository](#) by the author.
Items in Figshare are protected by copyright, with all rights reserved, unless otherwise indicated.

An innovative computer assisted proficiency test of English for academic purposes

PLEASE CITE THE PUBLISHED VERSION

PUBLISHER

© Loughborough University

LICENCE

CC BY-NC-ND 4.0

REPOSITORY RECORD

Aluisio, S.M., and Osvaldo N. Oliveira. 2019. "An Innovative Computer Assisted Proficiency Test of English for Academic Purposes". figshare. <https://hdl.handle.net/2134/1761>.

An Innovative Computer Assisted Proficiency Test of English for Academic Purposes

Sandra M. Aluísio
University of São Paulo (USP)
Department of Computer Science
CP 668, 13560-970
São Carlos, SP, BRAZIL
e-mail: sandra@icmc.sc.usp.br

Oswaldo N. Oliveira Jr.
University of São Paulo (USP)
Institute of Physics of Sao Carlos
CP 369, 13560-970
São Carlos, SP, BRAZIL
e-mail: chu@if.sc.usp.br

Abstract

A description is given of the **CAPTEAP** system that generates English proficiency tests (EPTs) automatically. Because the work is targeted at non-native, graduate students who need to develop technical writing skills, **CAPTEAP** tests the student's ability to understand and produce technical writing in English, also helping him/her develop the schemata for the academic discourse. For this reason the system was based on the Support Tool from the AMADEUS environment, which was developed for training students to write scientific papers. Computationally **CAPTEAP** was implemented on top of the Support Tool and the Hot Potatoes suite, thus allowing formative tests to be generated for researchers of Physics. When the package to generate summative tests is completed, we hope **CAPTEAP** to become a fully-fledged system for improving standardization and the accuracy of automatically generated EPTs. The advantage in comparison to conventional tests lies in the assessment of relevant skills graduate students really need throughout their academic career.

1. Introduction

In non-English speaking countries, graduate students are usually required to undertake an English proficiency test (EPT), aimed at assessing their ability to understand and produce technical literature, i.e. ability to grasp the message conveyed in a text passage and write in English. In practice, the test is performed by asking the student to translate a text passage from a magazine or technical book from English (L2) into L1 (the mother tongue), and in some cases produce a version in L2 of another piece of technical text originally in L1. It is usual, as in our university, that the test is prepared each year by a different member of the academic staff. This is disadvantageous since distinct exams may be highly non-uniform, owing to the high degree of subjectiveness in the evaluation process. There is a further limitation intrinsic in a type of exam in which only translations are requested: the academic communicative competence on the part of students is not evaluated in terms of the demands of highly standardized research articles written in English. Furthermore, the need of "genre-consciousness" is not aroused, which is essential for a novice researcher to better and faster perform the reading and writing tasks for his/her own research. As Hill et al. put it, "a knowledge of the rhetorical divisions of an experimental-research paper and the function of those divisions within the paper greatly enhances ESL (English as Second Language) student reading and writing skills" (Hill et al., 1982).

The purpose of this paper is to present the rationale employed in building the computer-assisted system **CAPTEAP** that fulfils the requirements of conventional proficiency tests, and takes the student one step further by **testing** his/her genre-consciousness, also **helping** him/her to develop the schemata for the academic discourse. The architecture of CAPTEAP encompasses both self-assessment tests and summative evaluation. A prototype system was built on top of two existing Windows softwares: a computer tool for training students to write scientific papers (Aluísio and Oliveira, 1995; Aluísio and Gantenbein, 1997a), named Support Tool, and the Hot Potatoes suite, a freeware package to create non-commercial educational Web materials developed by Half-Baked Software and the University of Victoria (<http://web.uvic.ca/hrd/halfbaked>).

The CAPTEAP architecture includes three modules — task creation, self-assessment and summative testing — which are plugged into the Support Tool allowing the same tool to be used for three different purposes. CAPTEAP is to provide a fully-fledged system for generating EPTs automatically, allowing adequate assessment of the relevant skills required by graduate students throughout their academic career. In Section 2, we dwell upon the motivation for designing and implementing CAPTEAP based on our previous experience in technical writing. Because CAPTEAP is somewhat merged with the Support Tool, the latter is briefly described in Section 3. The architecture of CAPTEAP, including samples of formative tests created for novice researchers of Physics, is presented in Section 4. Section 5 concludes the paper indicating further steps in the development of the whole system and how we plan to deliver it as a summative EPT in our university.

2. Background

We strongly advocate that English proficiency tests (EPT) for graduate students should be based on technical writing studies rather than only on tests of vocabulary, reading comprehension and grammar. A considerable bulk of literature exists for helping students in technical writing and reading, e.g. Trimble, 1985; Swales, 1990; Weissberg and Buker, 1990; Gosden, 1995, which was in fact used by the first author of this paper in 1998 in the EPT for the Computer Science Department of USP, São Carlos. In the test, one of the

exercises was designed to assess the students' consciousness about conventions in English usage in research papers. Guided (cued) exercises were designed, since they can help the student to get familiarized with the technical discourse. The set-up for the test was the following:

Goal to be met: how functional parts of a section of a research paper are arranged and how the linguistic forms match to rhetorical purposes.

Instructional material presented: a description of a schematic structure of an Introduction (i.e. pieces of information and order they are usually sequenced in research papers).

Support material: the title, Abstract and the Introduction section of a paper in their respective area: Computer Science, Statistics or Computational Mathematics.

The task: identify which sentences correspond to the parts/stages of the schematic structure.

The results were extremely encouraging because the students acknowledged that the test represented their first opportunity to analyze extracts of a scientific paper from the perspective of writing strategies. Even though 24 out of 37 students passed the exam, the overall performance in some exercises was poor especially because the students lacked experience in technical writing and in reading instructions carefully. For instance, several students included material from the Abstract in the segmentation exercise, despite the clear instruction to treat only the schematic structure of the Introduction. Most students failed to realize how much background information is usually provided in an Introduction before the Purpose of the paper is stated, particularly in interdisciplinary fields. Segments identified by the students were normally much longer than expected, for they failed to relate the various stages of an Introduction. These observations have stimulated us to try and implement a type of test where considerable information about writing strategies is provided to the student. For that we shall make use of writing tools we developed over the years, as explained later on.

The main shortcoming of the test is that it can only be applied by lecturers familiarized with technical writing strategies. This makes it impossible to apply the test on a regular basis with distinct lecturers. The way out was to replace the paper version of the EPT with summative tests for delivery on the Web. In order to do that we had to implement task bases associated with exercises from the technical writing literature and use computer-assisted assessment systems available in the market or a freeware from the WEB. One important pre-requisite was that the tool should be as slim as possible, in addition to providing its code open. This is why we did not choose software packages such as Question Mark for Windows (<http://www.questionmark.com/>), well recognized for online assessment, but rather selected the Hot Potatoes suite, a freeware developed specifically for delivering language learning activities. The latter is certainly much slimmer and seemed easier to build a prototype. The exercises created with Hot Potatoes use JavaScript for interactivity, and work under Netscape Navigator and Internet Explorer versions 3 and above, on both Windows and Macintosh platforms. The programs are designed so that almost every aspect of the pages can be customized through a graphical interface, but we intend also to adapt the HTML and JavaScript code to our needs. Our goal was to go through the first phase of the prototyping task: the creation of an abstract

layer on top of the basic types of exercise of Hot Potatoes, thus generating exercises specific to technical writing in English. In Section 4 the CAPTEAP system is described in detail.

3 An Overview of the Case-based Support Tool

The Support tool is part of a learning environment for scientific writing named AMADEUS (Fontana, 1993; Aluísio and Oliveira, 1995; Aluísio and Gantenbein, 1997b), which targets the understanding of the diverse processes involved in the creation of successful scientific papers. Its overall architecture embodies both advisory and tutoring tools. AMADEUS was motivated by previous research [Fontana-91] which used a framework of error categorization and error gravity analysis applied to theses, articles and research reports written in English by Brazilian graduate students. It was found that problems with inter-sentence relationships or with meanings of a more global nature, as well as with functional meanings in discourse, tend to be better treated by means of input material based on naturally occurring examples. In order to develop a writing tool to help non-native users, linguistic material was extracted from papers and books written by native users of English where reusable parts, usually rich in rhetorical contents, were highlighted (Aluísio and Oliveira, 1995). Such linguistic material offers microelements within a context, which form the macrostructure of the text framed by the schematic structure of its specific genre.

The reuse of linguistic material from a corpus, acquired either manually or semi-automatically, has been employed in a range of systems. These include report generators (Kukich, 1983, Smadja, 1991), hypertext-based support systems for software documentation (Born, 1992), case-based letter generators (Pautler, 1994), and document drafting tools (Branting, 1996, Paris, 1996). Within the AMADEUS Support tool, linguistic material is reused to improve the cohesion and coherence of a scientific paper's Introduction draft. The tool utilizes case-based reasoning (CBR) (Mantaras, 1995) to model the various stages of the writing process: planning, drafting and self-reviewing (Hayes, 1980). The detailed schematic structure shown in Figure 1 and the rhetorical structures of the Introduction corpus have been used in the Support Tool; one example is shown in Figure 2. These knowledge sources were applied to the three stages of the writing process. Accordingly, the user follows a three-step procedure: i) gathering of features, in which the user selects the features intended for his/her Introduction; ii) selection of the best-match case, following case recovery by the system; iii) revision/adaptation on the selected case in order to meet the user's purposes.

The screen-shots presented in Figures 1 and 2 illustrate the gathering and case recovery phases of the writing tool using a case base consisting of Physics papers. In Figure 1, the writer chooses the main components — setting, literature review, gap, purpose, methodology, main results, value of the research, and layout of the paper — and strategies — introducing research topic from the research area, historical review, listing criteria or conditions, for example — to be included in the introduction along with the writer's degree of certainty about their relative order in the text — sure about the order, some doubts, and many doubts.

To recover cases, three types of pattern matching between the selected features and the available cases are used: perfect match (equal lists), proper undermatch (sublist) and non-proper undermatch (intersection). The tool selects cases to be returned to the user by evaluating various combinations of these three similarity metrics, with the particular combination used depending on the degree of certainty of the user about the order of

components and strategies. Figure 2 shows one of the cases recovered from the components and strategies selected in Figure 1.

Figure 1: Gathering of structural features

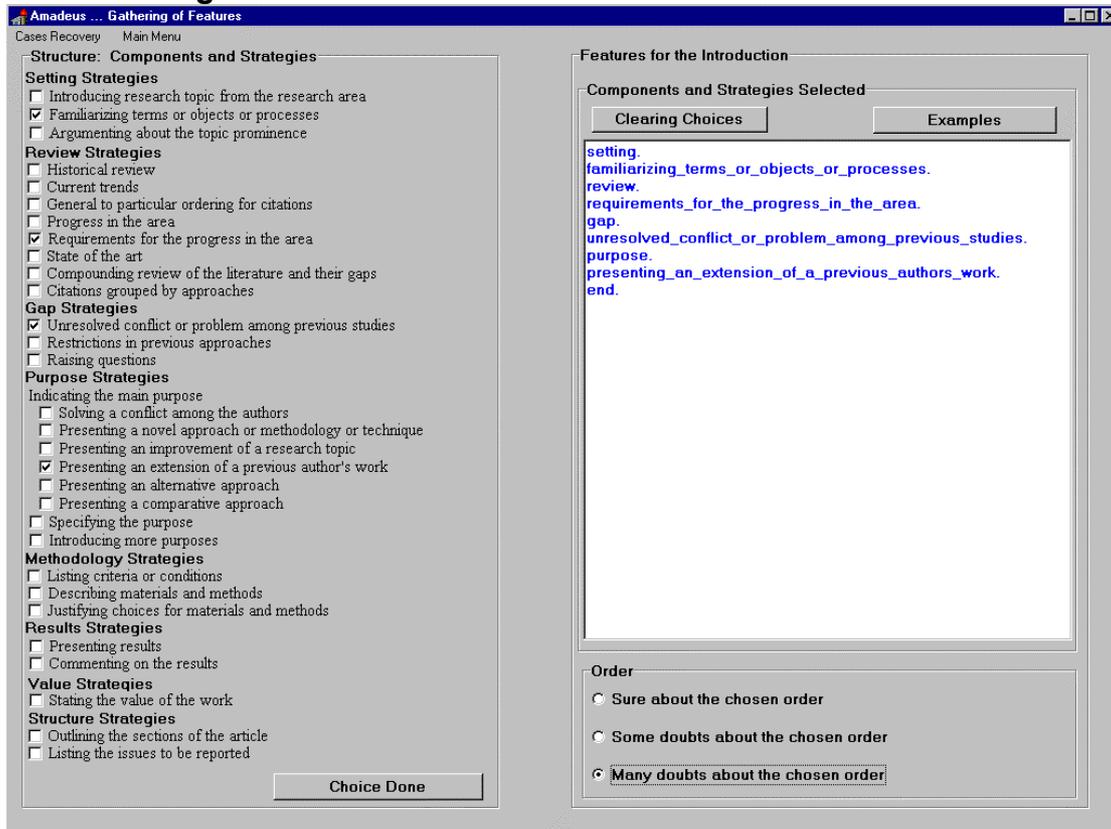
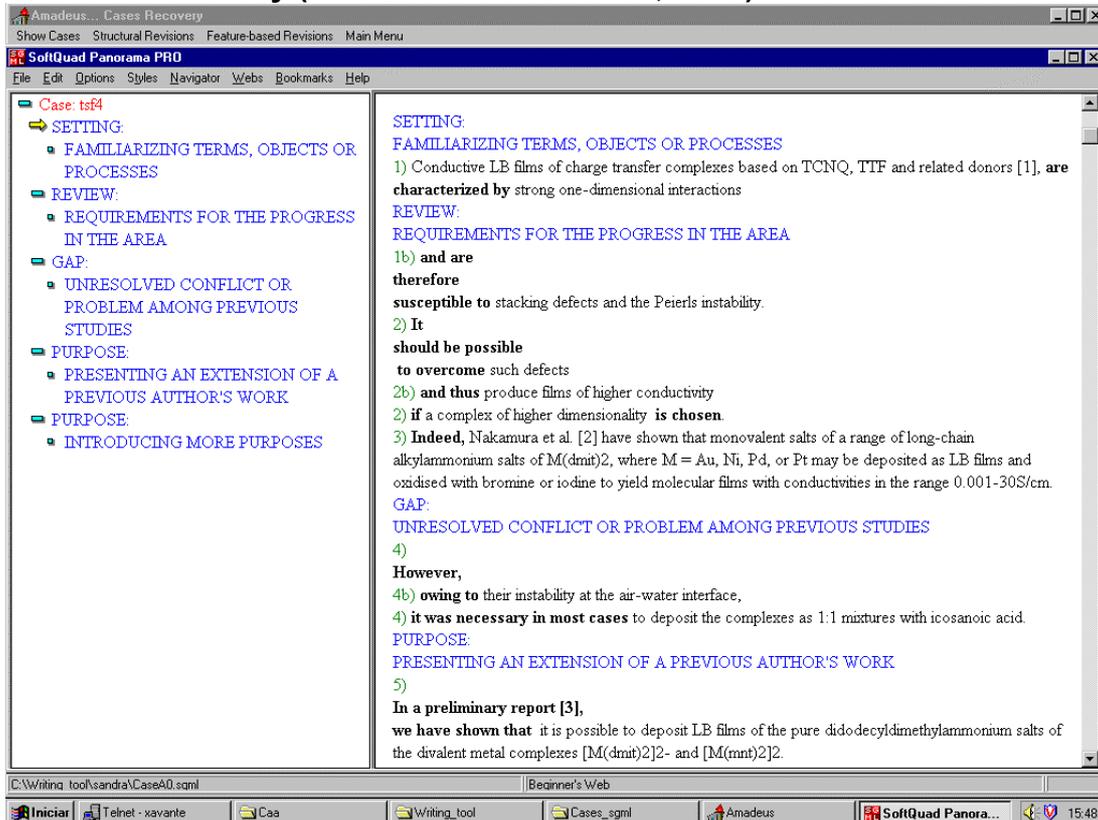


Figure 2: Case Recovery (from Aluísio & Oliveira, 1995)



According to Martin's approach to genre and register (Martin, 1992), the schematic structure of a genre constrains the combinatorial variable realizations of register, which in turn determine semantic/discourse and other linguistic structures. While many books on second language writing and professional reviewers focus mainly on lexico-grammatical and orthographical strata (for example, article use, lexical choice, syntactic structure, use of prepositions, and punctuation) in text writing, the organization of genres in terms of schematic structures also plays an important role (Ventola, 1991). Based on this perspective, Gosden (Gosden, 1995) outlines a method for analysis of textual revisions in scientific texts based on Systemic Functional Linguistics (SFL) (Halliday, 1993). The four categories of Gosden's revisions are:

1. addition of technical detail or statements **[+TD]**
2. deletion of technical detail or statements **[-TD]**
3. reshuffling of statements **[R]**
4. rhetorical machining **[RM]**

The last category (rhetorical machining) is subdivided according to three basic orientations:

- 4.1 rhetorical machining of discourse structure, by means of thematic (theme-rheme) and information (given-new) structures and cohesive relations **[RMd]**
- 4.2 changes which relate to writers' claims and writers' own research hypotheses and limitations of their work **[RMc]**, which also includes their research position to other published works
- 4.3 rhetorical machining which relates to the writers' purpose, reasons for, results of, research actions taken, and conclusions reached **[RMp]**

Our approach to self-review uses Gosden's four major categories of textual revision. In order to implement the revisions the corpus was analyzed to identify occurrences of features described in Table 1. For example, the features rhetorical *promotion* and rhetorical *demotion of technical detail* are related to the syntactic structure of sentences because there are three possibilities for clause/sentence combining: sentence connection/juxtaposition, clause coordination and clause subordination. Clauses can be promoted from a subordinate arrangement to a coordination or juxtaposition arrangement. The reverse operation is true for demotion. The features *markedness*, *hedging devices*, *reason*, *purpose* and *results* are related to the presence of different markers within the sentence. *Markedness*, for example, includes *in addition*, *here*, *furthermore*, as well as expressions of discourse organization, such as: *as shown in Figure 1*, *the first is*, *these are summarized*. Finally, the feature *modality* highlights the use of attitudinal disjuncts, i.e., adverbs used to comment on the content of the communication: *possibly*, *certainly*, *probably*, etc.

Table 1: Implementation of SFL-based self-review process

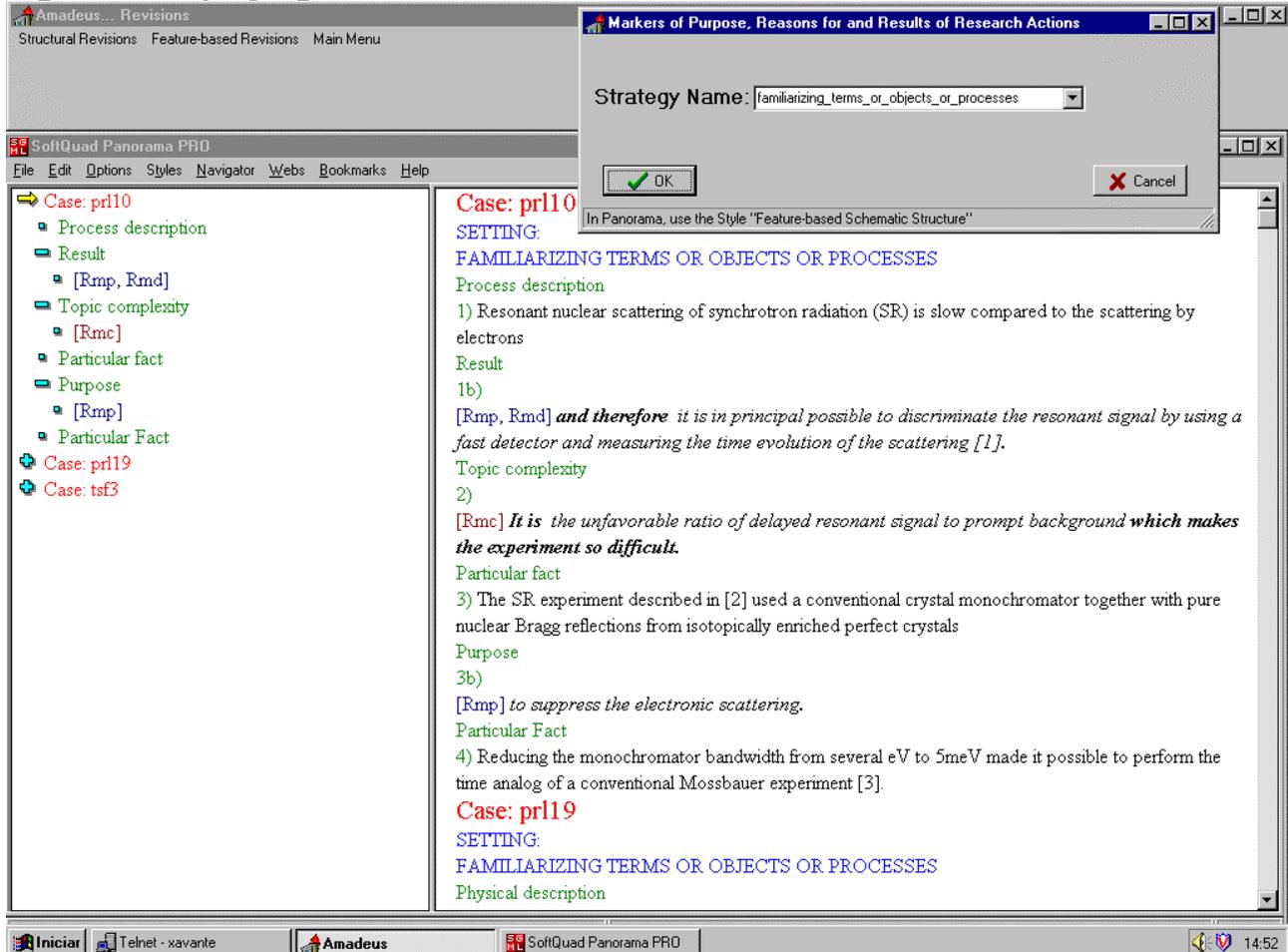
SFL-based revisions	Implementation
[-TD]	Deletion of optional messages in strategies
[+TD]	Addition of optional messages in strategies
[R]	Changing linguistic material and Exchanging a strategy/message for a similar one (features: rhetorical promotion and demotion of technical detail)
[RMd]	Changing linguistic material (feature: markedness)
[RMc]	Changing linguistic material (features: hedging devices and modality)
[RMp]	Changing linguistic material (features: reason, purpose, and result) and Addition of optional messages in strategies

The revisions receive names that are related to their effect on text rather than their method of implementation. They are activated according to the following nomenclature:

1. Addition of technical detail to illustrate ideas.
2. Deletion of technical detail when the paper addresses knowledgeable audiences or in short papers.
3. Promotion of technical detail to allow syntactic variety and differences in emphasis or importance.
4. Demotion of technical detail to allow syntactic variety and differences in emphasis or importance.
5. Addition of text organization markers (discourse markers) to link components, improving the information flow.
6. Addition of discourse markers to make the writer's claims about his/her research explicit.
7. Use of attitudinal disjuncts to comment on the content of the communication.
8. Addition of discourse markers to make the writer's research hypotheses or limitations explicit.
9. Addition of discourse markers to make the writer's research position in relation to other published work explicit.
10. Addition of markers to make the writer's purpose, the reasons for, and the results of, explicit.

The potential revisions on a text may be shown by using markers from Gosden's framework. Figure 3 illustrates one possible way to perform revisions using the text presented in Figure 2. In this example, bold text presents reusable material, italics highlights SFL-based revisions, while bold italics presents reusable material in the revisions.

Figure 3: Displaying the SFL-revisions on an Introduction



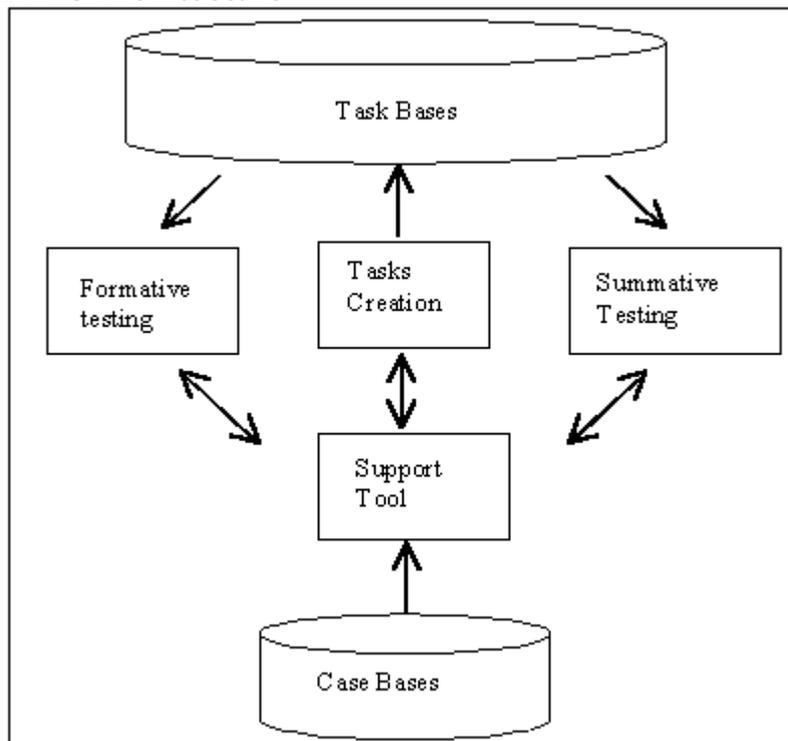
The natural synergy between CBR and SFL is an important factor in the effectiveness of the tool, since it follows the constructivist approach to learning - i.e. it emphasizes learning by doing. In addition, the philosophy underlying this approach highlights various ways of helping and supporting the students in order that they can be successful. For example: helping them learn to use cases effectively; giving them case libraries that will help them reason like experts; helping them learn the important areas to focus on; and helping them learn the variety of ways one goes about solving complex and open-ended problems such as the writing of texts.

There are, nevertheless, some drawbacks in using an educational tool, such as the Support Tool, in a non-supervised way. To facilitate non-supervised use, a novice student must be provided with a reflective learning module responsible for describing what skills are important to proficiency in a given field. This module will be explored in the following section, through description of the purpose and effects of each task proposed by CAPTEAP's formative tests.

4. Design and Description of CAPTEAP

Figure 4 shows CAPTEAP's architecture where three modules — task creation, self-assessment/formative and summative testing — are plugged in the Support Tool, with the plug action being shown in double arrows.

Figure 4: CAPTEAP's Architecture



Three software packages are created for different purposes. The first is to create task bases that will feed self-assessment and summative testing. The second is to allow reflective learning to the users of the Support tool that now possess a sequence of tests with real time feedback describing what skills are important to proficiency in technical writing. Finally, the third is to improve standardization of proficiency tests for academic purposes and to allow these tests to assess the knowledge and skill novice researchers need to better perform their activities in MSc and PhD programs. The resources of the Support tool will be available to the user during both self-assessment and summative tests because we hope to encourage students to develop the schemata for the academic discourse. In the summative mode the user will be allowed only one attempt to solve each task before submitting the answer. No hint or feedback button associated with the task being tested will be available, although he/she can access the resources of the Support Tool. In the self-assessment test, on the other hand, hint and feedback buttons are available, and many tries can be made until the user gets the correct answer.

We decided to design only guided (cued) exercises, since they can help the student to develop the language of the technical discourse. The first phase of our prototyping task was the creation of an abstract layer on top of the basic types of exercises of Hot Potatoes so that they could resemble exercises akin to the instruction of written English for science and technology. With this layer CAPTEAP works with 7 tasks, aimed at testing novice researchers':

1. genre-consciousness (ability to recognize the schemata for the academic discourse and the discourse marks related to the stages of the schemata). Two tasks are used to meet this target: **segmentation and reassembling of texts**. In the latter texts are presented in a scrambled form.

2. knowledge about reusable (conventional) expressions of scientific texts. The side effect is to test translation capabilities from L2 to L1. The task used is **completion of sentences**.
3. knowledge about English linguistic forms used to meet rhetorical purposes. The side purpose is to test translation capabilities from L1 to L2. We used a **reconstruction of sentences task**, with English sentences presented in a scrambled form.
4. ability to perform summarization operations. Two strategies to **summarizing texts** are used: reshuffling and deletion of technical details.
5. knowledge about rhetorical strategies (patterns of texts). We used an **association task** to match text to the respective pattern.
6. ability to reformulate text (expressions, clauses or sentences in a given pattern) to meet certain purposes. We used a **reformulation task** where the text, revisions operations and extra linguistic material is presented to users.

The tasks above allow 336 different types of exercises that may be automatically checked, since we can focus on 30 rhetorical strategies to do **association** exercises, and apply 10 revisions provided by the Support Tool to each rhetorical strategy when preparing **reformulation** exercises. The other tasks contribute with 1 (**segmentation, reassembling, completion, and reconstruction**) or 2 (**summarization**) types of exercises. An extra task could be used to test novice researchers' capability of free writing, although in our system it could only be checked manually (but see Burstein et al., 1998).

At the moment, CAPTEAP is a prototype working with two packages: one for generating tasks and the other for self-assessment. The task generator works as a wizard guiding the person responsible for preparing EST in the creation of tasks. It uses the resources from the Support Tool and five of the six tools of Hot Potatoes, since Jcross, which creates crosswords, has not been used. The resources from the Support Tool are:

- the schematic structure for research papers;
- the case base with 54 Introductions of Physics papers annotated in SGML;
- the case base of translated texts into Portuguese corresponding to the English papers, annotated in SGML;
- the case base written in Prolog which abstracts the rhetorical structure of each one of the 54 Introductions of Physics papers and code, also written in Prolog, to search for revisions on selected cases.

The tools employed from the Hot Potatoes suite are:

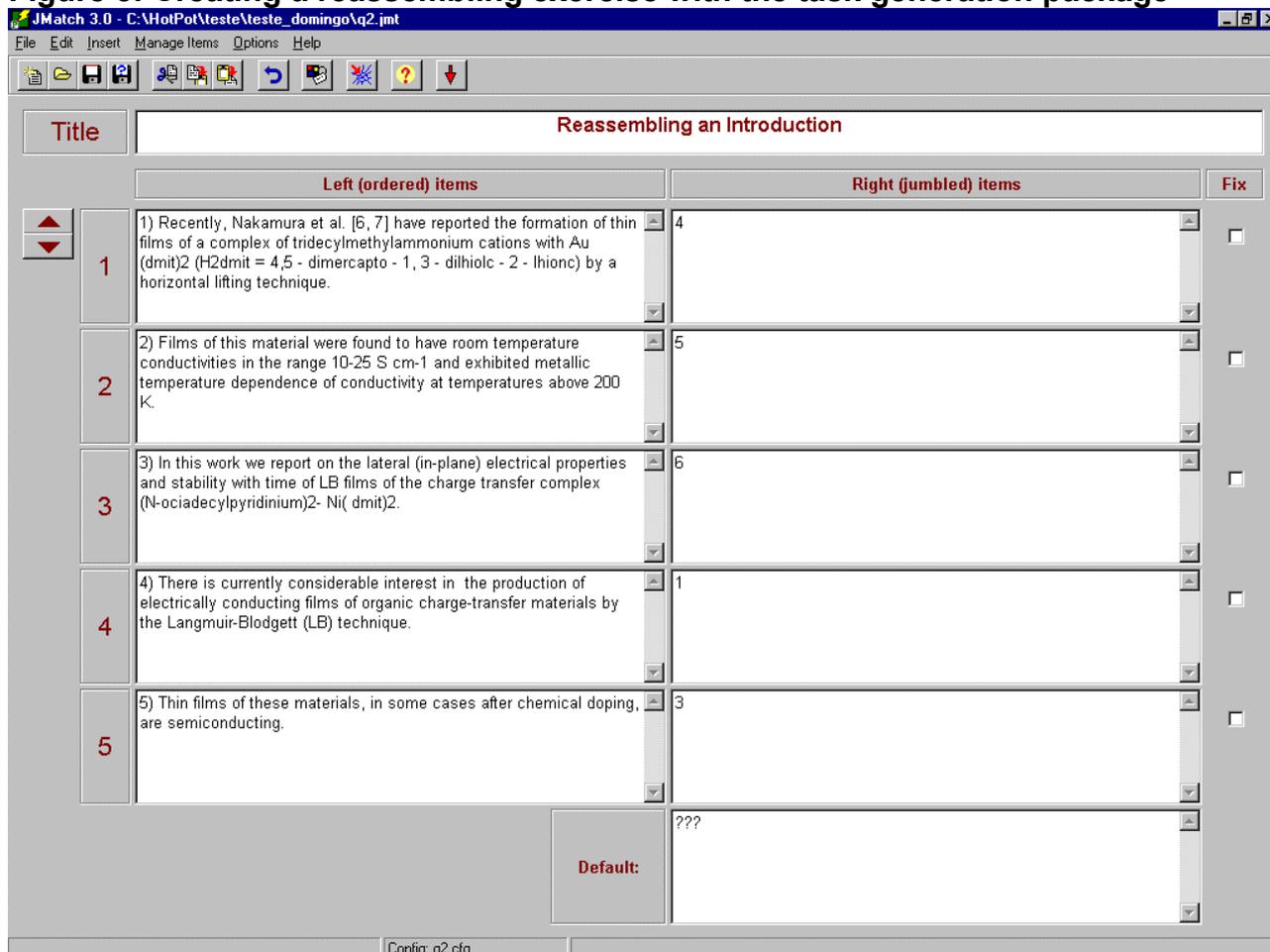
- JBC program (creates multiple-choice and true-false quizzes) to generate the segmentation and summarization tasks;
- Jcloze program (creates gap-fill exercises) to create the completion of sentences task;
- Jmix program (creates jumbled-sentence exercises) to create the reconstruction of sentences task;
- Jmatch program (creates matching and ordering exercises) for the reassembling and association task;
- JQuiz program (text-entry or short-answer quizzes) to create the reformulation task.

For example, in order to create an exercise for reassembling, the generation wizard guides the user:

(a) to select sentence by sentence from a case displayed in the browser SoftQuad's Panorama PRO (SGML browser: <http://www.softquad.com/>) as shown in Figure 2; (b) to

put them in the left side item column of Jmatch (Figure 5); (c) to copy their correct number in the right side column; and (d) to re-enumerate the left side column.

Figure 5: Creating a reassembling exercise with the task generation package



The self-assessment package, for formative testing, has added to the Support-Tool interface a pull-down menu with all the tasks available to allow the student perform several tasks. The links between the Support tool and the task being undertaken are made from feedback buttons and cued exercises that bring explanations for the tasks. For example, when the user is doing a reformulation task related to revision operations of addition or deletion of technical details, an explanation is presented of the possible revisions that can be performed on a strategy. The explanation is based on the notation for strategies used by the writing tool. The following strategy¹ from the Review component illustrates the kind of information the user may receive during the task:

Strategy: **Requirements for progress in the area**
[Topic Complexity]
Requirements for progress
Support

This can be interpreted as follows: the strategy *Requirements for progress in the area* is composed of the optional message *Topic Complexity*, followed by the message

¹ Strategies appearing in bold may be further refined and replaced by other strategies or rhetorical messages. Rhetorical messages appear in italics.

Requirements for progress and by the *Support* strategy appearing zero or more times. The *Support* strategy is defined as:
 (Motivation / Cause / Result / Purpose / Evidence / Particularization / Exemplification)

The formation rule for strategies makes the user aware of the kind of information a specific genre is supposed to contain. Moreover, optional messages and messages appearing zero or more times may be deleted or included if necessary, since they commonly carry details. We show in Figures 6 e 7 two samples of exercises for self-assessment, using the task base from Physics of CAPTEAP.

Figure 6: A segmentation exercise

An entire introduction from a paper in the field of Physics is reproduced here together with the description of summarization operations. First read them, then answer the questions.

Segmenting an Introduction

In the reading frame, an entire introduction from a paper in the field of Physics is reproduced together with the description of summarization operations. First read them, then answer the questions.

Help

1 Which sentence(s) in the introduction presented in the reading frame contains Stage I?

A 1
 B 2
 C 3
 D 1-4
 E none

2 Which sentence(s) in the introduction presented in the reading frame contains Stage II?

A 2-5

The structure of an Introduction

The introduction of a research paper can be divided into 8 parts, or stages.

In STAGE I, the writer establishes a context, or frame of reference to help readers understand how the research fits into a wider field of research.

In STAGE II, the writer reviews the findings of other researchers who have already published in his/her area of interest.

In STAGE III:

- * the writer indicates that previous literature described in stage II is inadequate because an important aspect of the research area has been ignored by other authors; or
- * the writer indicates that there is an unresolved conflict among authors of previous studies concerning the research topic; or
- * the writer indicates that an examination of the previous literature suggests an extension of the topic or raises a new research question not previously considered by other workers in the field.

STAGE IV formally announces the purpose(s) of the research.

STAGE V, the methodology, describes the steps followed in conducting the research.

In STAGE VI, results, the writer presents the main findings of the research.

STAGE VII indicates possible benefits or application of the research. Sometimes, stage VII is combined with stage IV in the same sentence.

Finally, in STAGE VIII, the writer presents the structure or layout of the paper, indicating the sections and briefly commenting on them.

Some of the stages are optional and may depend on the journal/conference style and length of the article. However, the general plan given in the reading

Figure 7: An exercise of the type “completion of sentences”

An introduction is given here paired with its translation in Portuguese. Try to complete the translation of Portuguese sentences based on the corresponding English sentences.

Translating L2 to L1

Fill in the blanks in order to complete the translation in Portuguese of the following English sentences:

Help

There is currently considerable interest in the production of electrically conducting films of organic charge-transfer materials by the Langmuir-Blodgett (LB) technique.

[] produção de filmes condutores de eletricidade a partir de materiais orgânicos de transferência de carga, [] Langmuir-Blodgett (LB).

Much of the work to date has been concerned with molecules such as tetrazalene (TTF) [1, 2] or tetracyano-p-quinodimethane (TCNQ) [3-5] which have been modified to make them amphiphilic.

[] moléculas [] o tetrazaleno (TTF) [1,2] ou o tetraciano-p-quinodimetano (TCNQ) [3-5], que foram modificadas para se tornarem anfifílicas.

Thin films of these materials, in some cases after chemical doping, are semiconducting.

Filmes finos desses materiais, em alguns casos após dopagem química, são semicondutores.

Recently, Nakamura et al [6, 7] have reported the formation of thin films of a complex of tridecylmethylammonium cations with Au (dmit)₂ (H₂dmit = 4,5 - dimercapto - 1,3 - dihiolc - 2 - lhionc) by a horizontal lifting technique.

[], Nakamura et al [6] [] a formação de filmes finos de um complexo de cátions de tridecil metil amônio com ouro (dmit)₂ (H₂dmit=4,5- di mercapto - 1,3 diol) através da técnica de deposição horizontal.

5 Concluding Remarks

In a first step, the CAPTEAP system and task bases will be made available for download with tests for Computer Science, Statistics, Computational Mathematics and Physics, in order to attend idiosyncrasies of the fields covered in graduate courses in our campus. We have already implemented two bases of Introduction sections: one with 54 annotated papers (cases) of Physics and another with 51 cases of Computer Science papers. There is a difference between the annotation scheme used for these bases which deserves some explanation. The Physics papers are from experimental research including sub-types of study, such as controlled experiments; correlational or case studies; or based on computer-generated models as described in Weissberg and Bucker (Weissberg and Bucker, 1990). Computer Science papers include types of study such as: empirical, experience, system, theory, methodology, as described in “CHI’96 Guide to successful Paper Submissions” (<http://www.acm.org/turing/sigs/sigchi/chi96/call/papers/papers-guide.html>). We have used a detailed schematic structure to annotate Physics papers, since our purpose was to help the writing of paragraphs. The annotation scheme of Computer Science papers was shallow in terms of schematic structure because we analyzed 5 types of study, as we meant to teach content information the studies refer to. The consequence of these different annotation schemes is the need to reformulate 3 tasks: summarization, reformulation of expressions, clauses or sentences, and association when working with Computer Science Papers or to go through a detailed analysis. The second requirement will be costly.

The aspects involving the design of the software package to create the secure session for summative decision making will be undertaken in a separate project. We are also considering to redesign the CAPTEAP prototype using software agents as in the CAETI project, aimed at combining tutors and existing software tools (Ritter and Koedinger, 1996). The task generation process should become automatic in the future. We may augment the interconnectivity between resources through messages sent to the student using the Support tool, as a way to identify possible misunderstandings and suggest the student to practise an exercise of a specific area of writing.

In conclusion, it is hoped that, in the long run, we may collaborate in the development of the last Generation of Large-Scale Educational Assessment — the Reinvention — described by Randy E. Bennett in (Bennett, 1998). In this generation, electronic learning tools for specific skills will be used for delivering instruction and certifying competence. In addition, we hope the assessment through **CAPTEAP** may improve standardization of proficiency tests, which would also be more accurate, and allow assessment of skills novice researchers really need.

Acknowledgements

The authors acknowledge the financial assistance from FAPESP and CNPq.

References

- Aluísio, S.M. and Oliveira Jr., O.N. (1995), A Case-Based Approach for Developing Writing Tools Aimed at Non-native English Users, Lecture Notes in Artificial Intelligence 1010, p. 121-132.
- Aluísio, S.M. and Gantenbein, R.E. (1997a), Towards the Application of Systemic Functional Linguistics in Writing Tools, Proceedings of International Conference on Computers and their Applications, pp. 181-185.
- Aluísio, S.M. and Gantenbein, R.E. (1997b), Educational Tools for Writing Scientific Papers. In: VIII Simpósio Brasileiro de Informática na Educação, ITA, pp. 239-253.
- Bennet, R.E. (1998), Reinventing Assessment. Technical Report, June, 1998. Princeton, NJ: Educational Testing Service. Available in <http://www.ets.org/research/pic/bennet.html>.
- Born, G. A (1992), Hypertext-Based Support Aid for Writing Software Documentation. In Computers and Writing - State of the Art, P. O'Brian-Holt and N.Williams (eds), Kluwer Academic Publishers, Dordrecht, pp. 266-277.
- Branting, L.K. and Lester, J. (1996), A Framework for Self-Explaining Legal Documents. In Proceedings of the Ninth International Conference on Legal Knowledge-Based Systems.
- Burstein, J., Kukich, K. Wolff, S. Lu, C. and Chodorow, M. (1998), Computer Analysis of Essays, NCME Symposium on Automated Scoring, April.
- Fontana, N. and Oliveira Jr., O.N. (1991) O Texto Acadêmico em Inglês como Língua Estrangeira - Dificuldades e Perspectivas. Atas do IX Simpósio Nacional de Ensino de Física, São Carlos, SP, pp. 571-576.

Fontana, N.; Aluísio Caldeira, S.M.; De Oliveira, M.C.F. and Oliveira Jr., O.N. (1993), Computer Assisted Writing--Applications to English as a Foreign Language. CALL, Volume 6(2), p. 145-161.

Gosden, H. (1995) Success in Research Article Writing and Revision: A Social-Constructionist Perspective. English for Specific Purposes, Vol. 14, No. 1, pp. 37-57.

Halliday, M.A.K. and Martin, J.R. (1993), Writing Science - Literacy and Discursive Power, University of Pittsburgh Press.

Hayes, J.R. and Flower, L.S. (1980), Identifying the Organization of Writing Processes. In L.W. Gregg & E. R. Steinberg (eds.), Cognitive Processes in Writing, Hillsdale, N.J. Erlbaum., pp. 3-30.

Hill et al., (1982) Teaching ESL students to read and write experimental research papers. TESOL Quarterly 16:333-47.

Kukich, K. (1983), Knowledge-Based Report Generation: A Knowledge Engineering Approach to Natural Language Report Generation. PhD Thesis, University of Pittsburgh.

Mantaras, R.L. and Plaza, E. (1995), Case-Based Reasoning. In The Newsletter of the European Network of Excellence in ML, pp. 29-37, Special Issue, September.

Martin, J. (1992), English Text: System and structure. Amsterdam: Benjamins.

Paris, C.L. and Van der Linden, K. (1996), Drafter: An Interactive Support Tool for Writing Multilingual Instructions. IEEE Computer, Special Issue on Interactive Language Processing.

Pautler, D. (1994), Planning and Learning in Domains Providing Little Feedback. In AAAI Fall Symposium on Planning and Learning Notes'94.

Ritter, S. and Koedinger, K. (1996), An Architecture for Plug-in Tutor Agents, Journal of Intelligence Artificial in Education, 7(3/4), 315-347.

Smadja, F. (1991), Retrieving Collocational Knowledge from Textual Corpora. An application: Language Generation. PhD Thesis, Computer Science Department, Columbia University.

Swales, J. (1990), Genre Analysis - English in Academic and Research Settings, Cambridge, UK: Cambridge University Press.

Trimble, L. (1985), English for science and technology: a discourse approach. Cambridge University Press.

Ventola, E. and Mauranen, A. (1991), Non-native writing and native revising of scientific articles. Functional and Systemics Linguistics - Approaches and Uses, Eija Ventola (ed.), Mouton de Gruyter.

Weissberg, R. and Buker, S. (1990), *Writing up Research - Experimental Research Report Writing for Students of English*. Prentice Hall Regents.