

## Reproducible models for CAT tools evaluation: A user-oriented perspective

Celia Rico  
Dpto. de Tecnologías de la Traducción  
Universidad Europea CEES, 28670 Madrid, Spain  
celia.rico@ti.fil.uem.es

*Abstract.* For some time now, we have been witnessing a gradual momentum which is remodelling the translation profession and transforming it from craft into an automated industrial process. In this context, the need for a coherent, reproducible model for translation tools evaluation emerges as a key element in the overall strategy for project management. This model should integrate the translation process into the evaluation procedure, considering at the same time stakeholders' needs and requirements.

The evaluation framework I propose is based on adopting a user-oriented perspective. This essential principle translates into defining, first, the context of use of the particular tool, and, then, into checking whether the system conforms to these specifications. The present paper rests on the preceding assumptions and is structured along the following lines:

- o *Scenario test outline.* This is broken down into (1) definition of stakeholders' profile as a set of characteristics covering such aspects as text type, translation environment, leveraging needs, team description and terminology requirements; and (2) definition of system performance criteria, using ISO 9126 as the starting point.
- o *Criteria weighting.* This implies assigning weighted values to the characteristics previously defined so as to decide their relative importance in the evaluation of the system performance.
- o *Metrics and Measurement.* A three element rating is defined for allocating scores to the system under evaluation.

## Introduction

Current trends in the language industry indicate that the new methods implemented in information management are pervading the translation world. In this respect, the field of information management is already directing its procedures towards managing interdependencies, since the complexity of producing documentation impels us "to abandon the silo perspective of product development, marketing communication, technical writing, translation, and product support" (Hofmann and Mehnert, 2000: 60). A translation project's life-cycle must comply, then, with the demands of a wider production environment which considers multilingual information management in the form of *information objects* (IO). An IO is a collection of information identified as a unit, and defined by its communicative purpose, the specific user it is addressed to, the business entity it represents (a product line or a corporate function), the information it provides (in a specific format and for a target audience) and some publishing restrictions (61).

One of the immediate consequences of this tendency is the high demand imposed on translation project management. This demand is higher than ever and introduces a challenge as new stakeholders enter the translation arena: clients, global teams of translators, cost, time, quality and resources, project's life-cycle, technology, and budget, among others. At the core of the whole process, and providing a means to face this challenge, we find Computer Assisted Translation (CAT) tools, a key resource with a constant presence in the translation workflow. From groundwork, term extraction and glossaries management, to text alignment and leverage of materials extracted from translation databases; from machine translation to post-editing; the operational environment of the translator is directly affected by technology.

In this context, the need for a coherent, reproducible model for translation tools evaluation emerges as a key element in the overall strategy for project management. This model should integrate the translation process into the evaluation procedure, considering at the same time stakeholders' needs and requirements. This should be a user-oriented framework for evaluation which could account for different translation scenarios: the industry, public administrations, agencies and freelancers.

Up to the present, many evaluations on translation tools have been conducted, specially in the case of translation databases (or translation memories) (see Benis, 1999; Assenat-Falcone, 2000) but, to my knowledge, these are not easily reproducible since they do not supply a set of parameters which could serve as a comprehensible model. In a sense, while these are certainly useful and valuable instances of evaluation, they concentrate on particular products used in a particular setting and under a specific environment which, in turn, restrains the system assessment to a predefined and specific set of needs. Acknowledging the benefits of such an evaluation, I believe that the current trends in the translation industry, as stated above, call for a broader model, one which can be easily reproduced and implemented according to different translation settings.

This paper is an attempt at presenting such a model. The point of departure is EAGLES<sup>1</sup> document EAG-EWG-PR.2 on *Evaluation of Natural Language Processing Systems* (1996) which outlines a comprehensive framework for evaluation and, in turn, can be adapted to particular, case-specific tools. In fact, this document advances a set of features for evaluation of translator's aids, concentrating on the description of user profiles, types of translator's aids and the functionality of translation memories.

The evaluation framework I propose is based on adopting a user-oriented perspective. This essential principle translates into defining, first, the context of use of the particular tool, and, then, into checking whether the system conforms to these specifications. This *golden rule* of evaluation implies the following: (a) designing an scenario test according to stakeholders' needs; (b) deciding how each feature of the scenario contributes to the final assessment of the system; and (c) executing the evaluation and assessing system performance. There is no such thing as a best system, but a best system for a particular situation.

---

<sup>1</sup> EAGLES's work (*Expert Advisory group for Language Engineering Standards*) is now directly linked to ISLE (*International Standards for Language Engineering*), a project for the development of standards in the Human Language Technology field. One of its working groups concentrates on evaluation. ISLE's website is available at: [http://www.ilc.pi.cnr.it/EAGLES96/isle/ISLE\\_Home\\_Page.htm](http://www.ilc.pi.cnr.it/EAGLES96/isle/ISLE_Home_Page.htm)

## Translation Scenario and Stakeholders

Any translation project, no matter how big or small, is defined, primarily, by two inseparable qualities: product and process. Translation is a *product* in the form of the final work the client is handed over, and it is defined at the same time as *process* as all the tasks that lead to this product. These two qualities conform the essential paradigm around which any *translation scenario* should be designed and directly affect the requirements and needs translation stakeholders may impose. The adequate identification of these forces, together with a correct estimation of their true contribution to satisfactory project completion, is the key to an objective evaluation of any translation tool. In fact, one of the main problems in evaluation is how to avoid subjective statements which, in any case, do not contribute to establishing relevant and efficient conclusions. While the core task in translation is apparently the same, methods vary, clients' expectations differ, teams of translators evolve and companies' profiles are different. Consequently, the evaluation of a particular tool in such a context and with so many forces interacting becomes a complex task which, in any case, should not lead to hasty results. Consequently, the set of features to be considered when designing the scenario test are distributed along four main categories: client, product, process and system.

### *Client*

A translation assignment is only originated when someone commissions it, imposing certain requirements on the job to be done. This means that there is no translation job as an *abstract entity* since there is always a client, usually more interested in the final product rather than on the process leading to it. In some instances what our client wants is speed and may even be willing to accept some trade-off with quality. It may also be the case that the customer places a strong interest in the means of communication with the translator, or the use of in-house specific terminology. No matter what the client's needs are, it is certainly true that listening to them and acting accordingly is a key to successful project completion.

In this sense, the client-translator relationship affects the definition of the translation scenario in the following aspects:

- o Compliance with client's conventions and style
- o Compliance with company's specific terminology
- o Preferred means of communication (for handing in assignments and returning finished jobs)
- o Deadlines
- o Average volume of translation
- o Required quality
- o Exchange of materials<sup>2</sup>
- o Pricing model

### *Product*

In general terms, we consider that the product of translation is the text. As a result of the advances in information technology and telecommunications, the text, as was traditionally conceived, has now evolved towards a typology previously unknown: Web sites, hypertexts, software to be localised, on-line product documentation and interactive guides, among others. The new translation products

---

<sup>2</sup> Schubert (1999) provides a detailed account of how sharing resources with your client affects the translation workflow and, hence, the choice of the tool. He identifies seven possible settings including, among others, a proprietary terminology database to be accessed on-line by the translator, a translation memory held by the translation provider to be used as a general memory for all translation tasks, or a proprietary translation memory held by the customer.

pose new challenges, both on the technical and the linguistic side, because content is now interactive, dynamic and subject to constant change, it is increasingly integrated in the final product and it has to be tailored to specific audiences in an unprecedented rush for reaching global markets.

While in translation theory there is no general agreement on a definition for the notion of text, it is still possible to identify three main characteristics: it is a coherent and cohesive unit, it is the product of a precise communicative action and the elements that conform the text are mutually relevant instead of a random collection of units (Cerezo, 2001). This definition accounts for all products the translation industry is currently dealing with, calling for new and different aspects to be considered for evaluation:

- o Format specifications
- o Content specifications: words, media elements, audio elements, maps, charts, video and animations, web links
- o Publishing needs
- o Reusability
- o Quality
- o Portability
- o Tailoring content for specific audiences
- o Linguistic consistency: spelling, date formats, referencing, style
- o Technical consistency
- o Front-end requirements
- o Back-end requirements
- o Languages involved

#### *Process*

The traditional appreciation of a translator's work environment has been that of a person who works alone, fenced behind a stack of dictionaries, paper and all sorts of documentation material. Not that this has ever been completely true, but it is certainly a picture surviving in the minds of some aliens to the profession. As a service, translation has customarily been postponed until the last phases of production, when all other aspects of product development have been completed. Translation has inevitably been tied up to the end of the process, either as an update of a manual or the manual itself, as a letter to a client, a best-seller to be introduced to a new cultural community, a contract establishing legal bounds between two parties or the marketing material of an international corporation.

Managing the translation process involves coordination, team work, planning and control techniques and these definitely affect modes of work and operational environments. On the one hand, there is no doubt the translation market is facing a dramatic increase as the figures demonstrate: the worldwide market for translation and localization is set to reach US\$20 billion in 2004, Microsoft executes over 1,000 localization projects a year with a revenue in fiscal 1998 of US 5\$ billion, there are practically 100,000 people translating professionally in Western Europe and 317,000 full- and part-time translators worldwide (Sprung, 2000: ix). On the other, the working environment of the translator is evolving towards decentralized global teams where technology is at the core of the process. In any case, it is certainly true that project management has lately gained a name in the translation profession due, mainly, to market growth and virtual teams. When translation is subcontracted to teams communicating through the Internet, productivity becomes the focus and it is in the area of planning, tracking and measuring for volume and quality where project management offers a wide experience (Devaux, 2000: 12).

In this complex paradigm and with so many forces interacting in the translation process, project management emerges as the key element marrying crafts, needs

and expertise. Translation teams share data, clients, targets, projects, people and resources, and it is the task of project managers to plan, instruct, monitor and control large amounts of data, quickly and accurately, while facilitating the problem-solving and decision-making process (Burke, 1999: 1). The features of the translation process that conform the translation environment can be stated, then, as follows:

- o Project size
- o Operational environment: technology to be used and company's resources, office space and equipments available
- o Timeframe
- o Budget and budget tracking
- o Future expansion
- o Quality assurance requirement (is ISO 9000 a requirement?)
- o Team available
- o Terms and conditions, as specified by the client, and subsequent assessment of their practicality
- o Training required and feasibility in terms of money, people and expectations
- o Outsourcing needs and logistic constraints
- o Viability of other translation projects running in parallel
- o Quotation by the word or by the hour. And if you charge by word, make sure you really count all words in the project!<sup>3</sup>
- o Management of updated
- o Testing results
- o Scalable work process

#### System

ISO 9126 standard ("Information Technology - Software product evaluation - Quality Characteristics and Guidelines for their Use) establishes six characteristics for software evaluation which provide an adequate framework for designing the translation scenario in terms of system characteristics. These six characteristics are summarised as follows<sup>4</sup>:

- a. Functionality: accuracy, suitability, interoperability, compliance and security
- b. Usability: understandability, learnability, operability
- c. Maintainability: analysability, changeability, stability, testability
- d. Reliability: maturity, fault tolerance, recoverability
- e. Efficiency: time behaviour, resource utilisation
- f. Portability; adaptability, installability, conformance, replaceability

In fact, the Expert Advisory Group on Language Engineering Standards (EAGLES) on its *Guidelines for Evaluation of Natural Processing Systems* adopts ISO terms and guidelines as its starting point and adds a seventh aspect, customizability, defined as "the ability to modify a product in order to satisfy a particular customer's needs" (Manzi et al, 1996).

<sup>3</sup> Counting words may turn a difficult, long task, since the *text* may contain graphic files, words in tags or in embedded scripts. Similarly, another aspect to consider is whether you are charging only translation or translation and verification of the running application. For an account of counting tools, see the section *Newcomers/Payment practices* at *TransRef* (<http://www.transref.org>). This is a portal for translators interested in translation technology, set up, maintained and updated with a professional mind.

<sup>4</sup> For a complete account of these characteristics see: <http://issco-www.unige.ch/projects/ewg96/ewg96.html>

For the purposes of this paper I will adapt EAGLES framework so as to present and adaptable model for system performance.

- o Functionality: this is broken down into:
  - Accuracy: measure system performance in terms of *precision* — percentage of valid segments from all those retrieved— and *recall* —percentage of segments retrieved from all those valid (for instance, when retrieving pairs of translated sentences from a TM).
  - Interoperability: check whether the system allows interaction with other translation aids.
  - Compliance with standards: check if the system supports different file formats.
  - Security: check whether the system covers your needs for translation validation and control of consistency over your team. If so, what are the mechanisms for translation validation?
- o Portability: if you need to reuse materials in other environments, does the system comply with the required standards?
- o Usability: how much effort is needed for recognizing the logical concept behind the system tasks and workflow? Usability also measures the effort needed for learning the application. What is the learning curve for effective use of the TM? What level of retraining does the TM impose on translation staff?
- o Efficiency: measure here time behavior in terms of retrieval time.
- o Maintainability: how does the system respond to fault tolerance and recoverability?
- o Back up and service: does the company selling the product offer a good service?
- o Pricing policy: can you afford to buy the system?
- o Investment in technical equipment: what level of investment is required before you can actually have the system running?
- o Customization: does the system allow easy customization?
- o Updates: is there any updating policy? Does it suit your needs?

Figure 1 illustrates a sample checklist with all the features contributing to the design of the translation scenario.

### **Criteria estimation**

Once the set of features that contribute to designing the translation scenario have been identified, the following step is to decide their relative value in the overall evaluation, i.e. assigning weighted values to each of the characteristics previously defined. This procedure rests on the assumption, as stated above, that different translation scenarios call for different translation needs and, therefore, different translation tools' requirements. Consider, for instance, how different team-management needs might be for freelance translators working on their own compared to those of an a translation department at an international corporation. The former would probably think that this feature is not applicable to their context of use, while the latter would assign a high relevance to this same characteristic.

Stakeholder	Feature	Weight	Score (100)
<b>Client</b>	<ul style="list-style-type: none"> <li>o Compliance with corporate conventions and style</li> <li>o Compliance with corporate terminology</li> <li>o Communication means</li> <li>o Deadlines</li> <li>o Average volume</li> <li>o Required quality</li> <li>o Exchange of materials (glossaries, translation databases)</li> <li>o Pricing model</li> </ul>		
<b>Product</b>	<ul style="list-style-type: none"> <li>o Format specifications</li> <li>o Content specifications</li> <li>o Publishing needs</li> <li>o Reusability</li> <li>o Quality</li> <li>o Portability</li> <li>o Tailoring content</li> <li>o Linguistic consistency</li> <li>o Technical consistency</li> <li>o Front-end</li> <li>o Back-end</li> <li>o Languages involved (does the system support them?)</li> </ul>		
<b>Process</b>	<ul style="list-style-type: none"> <li>o Project size</li> <li>o Operational environment</li> <li>o Timeframe</li> <li>o Budget</li> <li>o Future expansion</li> <li>o Quality assurance requirements</li> <li>o Translation team <ul style="list-style-type: none"> <li>• Teletranslation</li> <li>• Training required</li> <li>• Material sharing</li> </ul> </li> <li>o Reviewing needs</li> <li>o Quality control</li> <li>o Terms and conditions</li> <li>o Outsourcing needs</li> <li>o Running projects in parallel</li> <li>o Quotation</li> <li>o Management of updates</li> <li>o Testing results</li> <li>o Scalable work process</li> </ul>		
<b>System</b>	<ul style="list-style-type: none"> <li>o Accuracy (precision and recall)</li> <li>o Interoperability</li> <li>o Compliance with standards</li> <li>o Security</li> <li>o Portability</li> <li>o Usability</li> <li>o Efficiency</li> <li>o Maintainability</li> <li>o Backup and service</li> <li>o Pricing policy</li> <li>o Investment</li> <li>o Customization</li> <li>o Update policy</li> </ul>		

**Figure 1. The translation scenario: Sample evaluation checklist**

The model assumes that the ideal CAT tool scores a maximum of 100 points, this meaning the optimum performance. In order to find out how the different features contribute to the final score, the evaluator should go over each of them and assign a value according to the user's own needs. This value will later be weighted so that when all weighted values are added, they account for the expected optimum performance, i.e. 100 points<sup>5</sup>.

In order to illustrate how different scenarios are outlined in the evaluation model, figure 2 shows the values assigned to features under two different translation scenarios<sup>6</sup>.

*Scenario 1 (S1)* would correspond to the translation department of an international organization whose clients are, on the one hand, internal departments who produce confidential reports, economic reviews and summaries of conferences or papers. On the other hand, the department produces official documentation such as agreements, reports on technical issues or legally binding documents. Confidential and urgent translations are done in-house while lower priority documentation is outsourced.

*Scenario 2 (S2)* depicts translation management at a telecommunications company shipping their products and their corresponding literature and software simultaneously worldwide. The documentation is integrated in the product as an interactive guide but some printed material is also included. Translation requirements consist of software development, translation, localization and testing. In this sense, translation is part of product development and is always carried out in-house.

Comparing the values assigned to each of the features in the checklist, we see how differently those contribute to outlining the translation scenarios. For instance, the value assigned to *exchange of materials* is 0 for S1 and 100 for S2 to indicate that in S1 the possible exchange of translation materials is not an essential feature for a CAT tool evaluation, while in S2 this is a key characteristic. Conversely, the values given to *outsourcing needs* indicate, for S2, that all translation is done in-house, whereas for S1, the assigned value (30 points) means that some 30% of the translation work is done externally.

## Metrics and Measurement

In their evaluation of two dictation systems, Canelli *et al* (2000) applied a quality model that rested on ISO/EAGLES methodology and developed a three element rating scale used to indicate whether the metric applied counted as *good*, *acceptable* or *unacceptable*. Each metric would then produce a score and "the number of points awarded to the product is determined by its position on the three point scale, so that if the score is good, the total number of points is given, if acceptable, half of the points, if poor, none at all".

---

<sup>5</sup> To calculate the weighted values, multiply each value (V) by a constant (C) defined as  $C = \frac{W}{\sum V}$

where C is the constant, W is the optimum performance (100) and V is the value assigned to each feature.

<sup>6</sup> Both scenarios are simulated examples.



Stakeholder	Feature	Scenario 1		Scenario 2	
		Av	Wv	Av	Wy
<b>Client</b>					
	o Compliance with corporate conventions and style	100	2,92	100	2,14
	o Compliance with corporate terminology	100	2,92	100	2,14
	o Communication means	50	1,46	100	2,14
	o Deadlines	75	2,19	100	2,14
	o Average volume	100	2,92	100	2,14
	o Required quality	100	2,92	100	2,14
	o Exchange of materials	0	0	100	2,14
o Pricing model	0	0	100	2,14	
<b>Product</b>	o Format specifications	100	2,92	100	2,14
	o Content specifications	100	2,92	100	2,14
	o Publishing needs	10	0,29	100	2,14
	o Reusability	5	0,15	100	2,14
	o Quality	100	2,92	100	2,14
	o Portability	0	0	100	2,14
	o Tailoring content	0	0	100	2,14
	o Linguistic consistency	100	2,92	100	2,14
	o Technical consistency	100	2,92	100	2,14
	o Front-end	0	0	100	2,14
	o Back-end	0	0	100	2,14
	o Languages involved	100	2,92	100	2,14
<b>Process</b>	o Project size	100	2,92	100	2,14
	o Operational environment	80	2,34	100	2,14
	o Timeframe	100	2,92	100	2,14
	o Budget	50	1,46	100	2,14
	o Future expansion	0	0	100	2,14
	o Quality assurance requirements	0	0	0	0
	o Translation team;				
	Teletranslation	0	0	0	0
	Training required	0	0	0	0
	Material sharing	0	0	100	2,14
	o Reviewing needs	0	100	100	2,14
	o Quality control	100	2,92	100	2,14
	o Terms and conditions	0	0	100	2,14
	o Outsourcing needs	30	0,88	0	0
	o Running projects in parallel	0	0	100	2,14
	o Quotation	0	0	100	2,14
	o Management of updates	0	0	100	2,14
	o Testing results	0	0	100	2,14
o Scalable work process	20	0,58	100	2,14	
<b>System</b>	o Accuracy (precision and recall)	100	2,92	100	2,14
	o Interoperability	100	2,92	100	2,14
	o Compliance with standards	20	0,58	100	2,14
	o Portability	20	0,58	100	2,14
	o Usability	100	2,92	100	2,14
	o Efficiency	100	2,92	100	2,14
	o Maintainability	100	2,92	100	2,14
	o Backup and service	100	2,92	100	2,14
	o Pricing policy	100	2,92	80	1,71
	o Investment	20	0,58	100	2,14
	o Customization	100	2,92	100	2,14
	o Update policy	80	2,14	100	2,14

\*Av - Absolute value

Wv - Weighted value

**Figure 2: A checklist for two different translation scenarios**

Similarly, the present model for CAT tools evaluation assumes that once the translation scenario has been defined as a set of characteristics and all their corresponding weights have been established, the actual analysis of the system will produce a score for each of the features, ranging from the assignment of the maximum weighted value of an attribute when the performance is optimum, through half this value for partially satisfactory performance, to zero for bad or unsatisfactory performance.

In executing the evaluation, the evaluator should proceed by checking, in turn, all the features in the checklist and then allocating scores according to the following metrics:

1. Optimum performance: allocate maximum score as stated in the weighted value.
2. Partially satisfactory performance: give half the possible score.
3. Poor performance: allocate no points at all (0).

When all scores have been assigned, the evaluator adds them up and, for the purposes of the evaluation report, will consider that the best system scored closer to 100 points.

## Discussion

Following King's (1999) *recipe* for designing an evaluation framework for machine translation we find that "the ultimate question is whether the object being evaluated fits what the customer of the evaluation wants or needs". The present paper outlines the means to answer this question and, in order to show the practicalities of the model proposed, I will now check it against the seven steps for evaluation as stated in King.

### *Step 1. Why is the evaluation being done?*

This is a stage prior to the actual evaluation but still an essential one. The key is to find the purpose of the evaluation so that from the very beginning the user and the evaluator work on common grounds. At this point, we should consider why we are to evaluate a particular tool: is it for the sake of speed in translation?, for an increase in quality?, because our client is supplying us with certain materials in a certain format?, or because we want to set up our own workstation?

### *Step 2. Elaborate a task model*

Our task model is defined by the translation scenario where the tool under evaluation will be used.

### *Step 3. Define top level quality characteristics*

These are defined as the set of translation characteristics corresponding to the four translation stakeholders, i.e. client, product, process and system. Since not all attributes are of equal importance each of the characteristics is assigned a weighted value, as shown above.

### *Step 4. Produce detailed requirements*

In King's words, "a valid and reliable way of measuring how a system fares with respect to the attributes must be found". Nevertheless, finding attributes which are measurable is sometimes impossible as many of the characteristics in the translation scenario are not numerically quantifiable. This is where the evaluation framework introduces some unavoidable subjectivity which I try to compensate by defining a three element rating

scale to account for excellent performance, partially satisfactory performance and poor performance.

#### Step 5. Define metrics

Metrics and measurement in the model are conceived with the challenging task of converting into a numerical score some observed characteristics which are, otherwise, qualitative. As is the case in many evaluation methodologies (especially in evaluation of machine translation systems), finding agreement across a representative community of users on what counts as good or poor performance is normally difficult (King, 1996). In this case, metrics are based on "internal validity" rather than on "external validity", which means that "their interpretation is constrained only by the individual's notion of what the metric means" (193). At best, agreement can be negotiated among users under the same translation scenario.

Steps 6 and 7. Design the execution of evaluation and Execute evaluation  
So far, the model presented here has been tested in simulated scenarios and a through exploration of its practicality under a real translation setting is the subject for further work.

## Conclusion

The evaluation model presented in this paper is inspired by a user-oriented perspective which could account for different translation scenarios such as the industry, public administrations, agencies and freelancers, and their different translation needs. The model acknowledges these differences and brings up a series of characteristics that correspond to the four main actors in translation, i.e. the client, the product, the process and the CAT tool. Under this framework, the translation characteristics that conform the scenario are easily extensible to allow for other aspects not originally considered because the process of assigning weights and estimating values remains unchanged.

## References

- Assenat-Falcone, S. (2000): "More Translation Memory Tools (not many more, but good ones)" in *Translation Journal*.  
Available at: <http://www accurapid.com/journal/12TM.htm>
- Benis, M. (1999): "Translation Memory from O to R" in *TransRef*, available at <http://www.transref.org/u-articles/Benis3.asp>. Originally published in ITI Bulletin.
- Burke, R. (1999): *Project Management. Planning and Control Techniques*. New York: John Wiley.
- Canelli, M., D. Grasso and M. King (2000): "Methods and Metrics for the Evaluation of Dictation Systems: A Case Study", in *Second International Conference on Language Resources and Evaluation (LREC)*. Athens, Greece, vol. II: 1325-1331. Available at: <http://www.lrec-conf.org/lrec2000/www.xanthi.ilsp.gr/lrec/conference/papers-14.htm>
- Cerezo Ceballos, L. and G. Corpas Pastor (2001): "Estrategias para no naufragar en la traducción de hipertextos", in *El traductor profesional ante el próximo milenio*. Madrid: Universidad Europea CEES.
- Devaux, S. A. (2000): "Getting a Grip" in *Language International*, April 2000: 12-13.
- EAGLES (1996): *Evaluation of Natural Language Processing Systems*, Eagles document EAG-EWG-PR.2.  
Available at: <http://issco-www.unige.ch/projects/ewg96/ewg96.html>

- Hofmann, C. and T. Mehnert (2000): "Multilingual Information Management at Schneider Automation", in Sprung, R. C. (ed.): *Translating into Success. Cutting-edge strategies for going multilingual in a global age.* Amsterdam/Philadelphia: John Benjamins.
- King, M. (1996): "On the Notion of Validity and the Evaluation of Machine Translation Systems", in Somers, H. (ed.): *Terminology, LSP and Translation.* Amsterdam/Philadelphia: John Benjamins.
- King, M. *et al* (1999): "MT Evaluation", in *MT Summit VII*, Singapore, September 1999. Available at:  
<http://www.issco.unige.ch/publications/publication-list.html#1999>
- Manzi, S., M. King and S. Douglas (1996): "Working Towards User-oriented Evaluation" in *Proceedings of the International Conference NLP+IA/TAL+AI*, Moncton, NB, Canada: 155-160. Available at:  
<http://www.issco.unige.ch/publications/publication-list.htm#1996>
- Schubert, K. (1999): "Resource and Workflow Management Support in Teletranslation", in *Translating and the Computer 21*, London: Aslib.
- Sprung, R.C. (2000): *Translating into Success.* Amsterdam/Philadelphia: John Benjamins.

## **Author**

Celia Rico, Ph.D., is the head of the Department of Translation Technologies at the Universidad Europea CEES (Madrid, Spain), where she trains future translators on the fundamentals of CAT tools as well as on the principles of translation strategies. Dr. Rico's publications have concentrated on areas such as translation memory evaluation, machine translation and translator competence, and the impact of new technologies on the translation profession.

Dr. Rico can be reached at [celia.rico@ti.fil.uem.es](mailto:celia.rico@ti.fil.uem.es)