

A lo largo de sus cincuenta años de éxitos y fracasos, la investigación en Traducción Automática ha seguido caminos divergentes, reinterpretando el papel de la máquina y del humano en el proceso de automatización. Desde una perspectiva pedagógica, la TA ha atraído—y atrae—a estudiantes procedentes de una amplia gama de disciplinas. Esto plantea un gran reto a los formadores, que se ven obligados a adaptar sus programaciones docentes en función de las destrezas particulares de cada comunidad meta. El presente trabajo aborda algunas de las cuestiones clave en la enseñanza de la traducción automática. La primera sección repasa la noción de TA, reubicándola en un marco de referencia más amplio. La segunda presenta una taxonomía de potenciales discentes. Las siguientes secciones (3, 4 y 5) se centran en las principales comunidades meta de la enseñanza de TA en la actualidad (estudiantes de Lingüística Computacional, Traducción y Lenguas Extranjeras, respectivamente). La sección sexta presenta brevemente nuestras conclusiones.

# Teaching about Computers and Translation

LUIS CEREZO CEBALLOS  
*Georgetown University*

*Along 50 years of successes and failures, research into Machine Translation has evolved in divergent directions, reinterpreting the role of machines and humans in the automation process. From a pedagogical perspective, MT has attracted—and attracts—students from very different backgrounds. This poses important challenges for instructors, who have to adapt their curricula according to each target group, weighing their assets and weaknesses. This paper aims to bring forth into the arena some of the questions behind teaching Machine Translation. Section 1 revisits the notion of MT, locating it within a general frame of reference. Section 2 presents a taxonomy of potential students. The following sections (3, 4, 5) focus on today's main target groups in the teaching of MT (CL students, trainee translators, and FL learners, respectively). Section 6 is a brief conclusion.*



## 1 MT REVISITED: A HISTORICAL FRAME OF REFERENCE

12

During its 50 years of history, research into MT has evolved in different directions<sup>1</sup>. In the early years (1950s), the prevailing assumption among researchers was that MT should achieve results comparable to those of human translators. As a result, they focused on the intersection of two objects of study, viz. *translation* and *machines* (Fig. 1).

In a second stage of research (1960s), the previous assumption was considered too ambitious. It was first criticized by Bar-Hillel at MIT, who suggested that semantic barriers could only be overcome by feeding computers with knowledge about the real world. More categorically, the 1966 ALPAC report predicted little prospect of useful MT in a foreseeable future and recommended instead the development of machine aids for translators. The *human* component was thus incorporated into the previous approach (Fig. 2).

As a result of the ALPAC report, research on MT was clearly forked into two different branches: (i) Machine Translation as such (MT)<sup>2</sup> vs. (ii) Machine Aided Translation (MAT)<sup>3</sup>. This double scenario has remained valid ever since. Thus today's MT researchers can be divided into two main categories: those in favour of an *exclusive* approach (MT

<sup>1</sup> Material on the history of MT is based on Hutchins and Somers (1992: 5-9).

<sup>2</sup> MT is often referred to as Fully Automatic Translation (FAT) (Sager, 1994: 290) or Fully Automatic High Quality Translation (FAHQT) (Hutchins and Somers, 1992: 147).

<sup>3</sup> Many authors offer a more fine-grained perspective of MAT, defining it as the merging of two approaches, viz. Human-Aided Machine Translation (HAMT) and Machine-Aided Human Translation (MAHT) (cf. Hutchins and Somers, *op. cit.*; Sager, *op. cit.*).

<sup>4</sup> This is in fact the title of a forthcoming collective monography edited by Somers (see *References*).

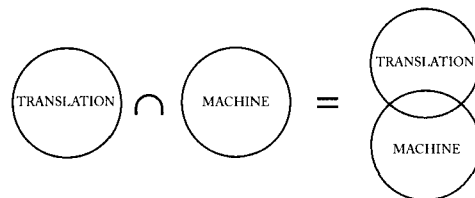


Fig. 1: Main objects of study during the first stage of research into MT

research should focus exclusively on the study of standalone translation software) and those in favour of an *inclusive* approach (MAT tools should also be taken into consideration).

One representative of the latter approach is the European Association for Machine Translation. The translation software now available —argues EAMT (2001)— goes beyond standalone translation programs and runs the gamut «from simple dictionary lookup programs used as word-processor add-ons to soph[is]ticated batch-translation systems based on relational databases running under Unix». For this reason, *Machine Translation* —defined as «the application of computers to the task of translating texts from one natural language to another»— is nowadays an «archaic-sounding» term. In an attempt to provide solutions to this terminological problem, some authors have subsumed both MT and MAT under a more general heading, e.g. *Computers and Translation*<sup>4</sup> (Fig. 3).

## 2 TEACHING ABOUT COMPUTERS AND TRANSLATION: A TAXONOMY OF POTENTIAL ADDRESSEES

As stated by Forcada et al. (2001: 3), the cross-disciplinary and multi-directional nature of M(A)T poses «important challenges» on

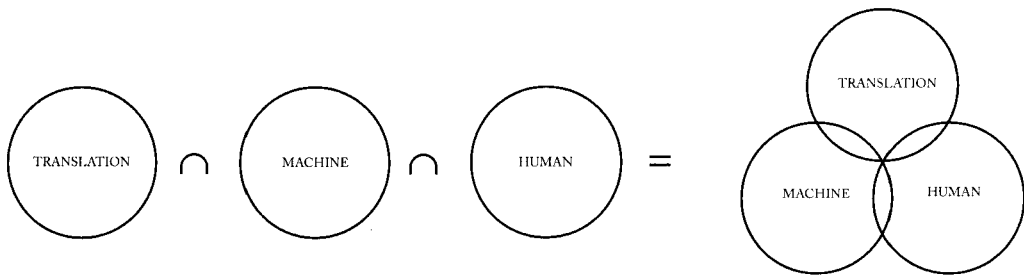


Fig. 2: Transition into Machine Aided Translation

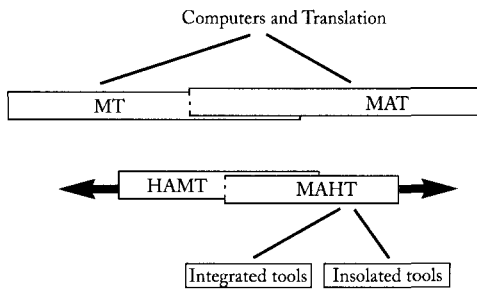
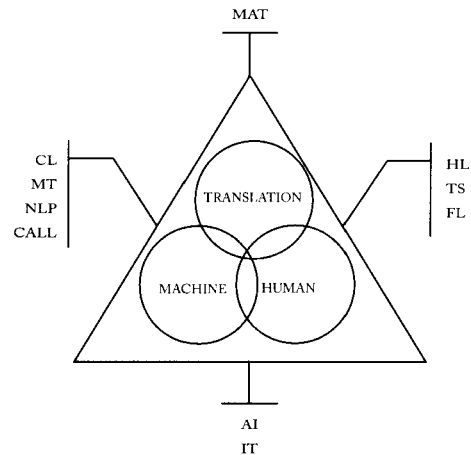


Fig. 3: Relocation of the notion of MT (Hahn)

**Key:**

CL	COMPUTATIONAL LINGUISTICS
MT	MACHINE TRANSLATION
NLP	NATURAL LANGUAGE PROCESSING
CALL	COMPUTER-AIDED LANGUAGE LEARNING
MAT	MACHINE-AIDED TRANSLATION
HL	HUMAN LINGUISTICS
TS	TRANSLATION STUDIES
FL	FOREIGN LANGUAGES
AI	ARTIFICIAL INTELLIGENCE
IT	INFORMATION TECHNOLOGIES

Fig. 4: Domains related to the intersection of sets translation, machine, and human

instructors, who have to deal with students from very different backgrounds. This multi-directional aspect results from the intersections of the three spheres discussed above—*translation, machine, human* (Fig. 4).

As shown in the figure, four intersection areas can be distinguished:

(i)  $TRANSLATION \cap MACHINE$ : this area is predominantly the object of study of computational approaches to translation, viz. Computational Linguistics (CL), Machine Translation (MT), and Natural Language Processing (NLP). However, it has been suggested that some peripheral disciplines, such as Computer-



Aided Language Learning (CALL), could also benefit from studying this intersection area. This does not seem foolish, since translation is a consolidated practice in L2 acquisition.

(ii) TRANSLATION  $\cap$  MACHINE  $\cap$  HUMAN: this area is covered by computational approaches to translation with a view to human interaction, viz. Machine Aided Translation (MAT).

(iii) TRANSLATION  $\cap$  HUMAN: this is the area of non-computational approaches to translation, running all the gamut of *relevant* disciplines traditionally subsumed under General Linguistics—hence our proposal Human Linguistics (HL), as opposed to Computational Linguistics (CL). This is also the place for Translation Studies (TS), Translation theory, or so-called Translatology. Lastly, this area is often studied in the teaching of Foreign Languages (FL), since as mentioned above, translation exercises are widely used in L2 acquisition.

(iv) MACHINE  $\cap$  HUMAN: this intersection area corresponds to disciplines such as Artificial Intelligence (AI) and Information Technology (IT).

This paper focuses on how to teach about computers and translation to students and professionals with an interest in any of the first two groups of disciplines. The task now is to classify these target students from a real market perspective, with the aim of designing curricula that adequately meet their needs. According to the majority of academic teachers, two target communities can be distinguished: trainee translators and students of computational linguistics. A third community is suggested by a more restricted group of academics—students of foreign languages. As Somers puts it (2001: 25):

The use of MT and related software in the classroom is motivated by different concerns:

one relates to teaching about computers and translation for its own sake, as part of course in one of the contributing fields such as linguistics, computational linguistics, computer science, information technology and so on. Another is teaching trainee translators and other professional linguists about translation software. A third is the role (if any) of this software for teaching languages.

In the next sections we will discuss the topic of teaching M(A)T from this three-fold perspective. But first we would like to concentrate on some preliminary considerations.

In contrast with the vast number of publications on general aspects of MT, literature on pedagogical aspects is quite restricted and very recent (cf. Somers, 2001: 25). However, the advent of the Internet has opened many doors. As Balkan (2001: 7) remarks, the Web provides «fast, usually free, access to **reference material**<sup>5</sup> which might otherwise be difficult and/or expensive to track down.» In addition to that, the Web is an unrivalled pool of **MT resources**. As Forcada et al. (2001: 3) state, «due to the growth of the internet, both commercial and experimental machine translation systems are more readily available than ever». At a third level, the Internet is seen as a powerful **teaching medium**, since «it may naturally integrate real MT systems as part of the learning environment» (Forcada et al., 2001: 3). This is interpreted by Balkan (2001: 7) as an open door to **distance learning**, as online access «makes materials available to a wider body of students, since distance is no barrier». On the negative side —adds Balkan (2001: 9)— the **scarcity of information** about many of the MT systems available online «limits their usefulness as teaching aids». To tackle this problem, Balkan (2001: 11) suggests creating an **MT portal**, which would ideally contain «not just a list of resources, but one that has been properly annotated and preferably indexed to allow for easy browsing and searching. Retrieval could be on

<sup>5</sup> (Bold ours).

## DOMAIN

(i) COMPUTER PROGRAMMING

(ii) LANGUAGE SKILLS

(iii) FORMAL LINGUISTICS

(iv) NLP

## SUBDOMAIN

Procedural programming (Java)  
 Declarative programming (Prolog)  
 Perl for corpus manipulation  
 CGI programming  
 Algorithms and Data Structures  
 Computability and Complexity

French, German or Spanish  
 Translation and Culture and Society

Phonetics and Phonology  
 Syntax and Semantics  
 Morphology

Logic, Maths, and Statistics  
 Corpus Linguistics  
 Artificial Intelligence  
 Parsing and Statistical NLP



Table 1: Academic curriculum of ACL at DCU (Kenny and Way, 2001: 14)

the basis of language, topic (e.g. statistical MT), type (e.g. software tool, book, etc.), and level (e.g. technical, non-technical)».

### 3 TEACHING TO STUDENTS OF CL

As stated by Balkan (2001: 8), a student of CL «is likely to be well-versed in both computers and linguistics, and therefore able to handle complex descriptions of systems and translation engines». Ideally, the academic background of a student of CL should include both theoretical and practical training in four different domains, namely (i) computer programming, (ii) foreign language skills, (iii) formal linguistics, and (iv) NLP. For illustrative purposes we have selected the academic curricu-

lum of the degree in Applied Computational Linguistics (ACL) at DCU (Kenny and Way, 2001: 14) (Table 1).

As opposed to trainee translators and FL learners, students of CL will have a primary focus on MT systems rather than on MAT applications, and will prefer experimental systems to commercial ones, as the working mechanisms of the latter are rarely made explicit, for proprietary reasons.

According to Clavier and Poudat (2001: 22) MT courses play a crucial role for students of CL, since MT systems put into practice the theoretical contents of all the domains they have received tuition in. As Somers (2001: 26) puts it<sup>6</sup>:

<sup>6</sup> All bold and italics are ours.



For the student (and teacher) of CL, then, MT systems can be used to illustrate problems (and solutions) in language analysis at various levels both monolingually and contrastively. **Source-text analysis** requires morphological disambiguation and interpretation, word-sense disambiguation, syntactic, semantic and pragmatic disambiguation. **Translation** involves converting linguistic aspects of the source text into their appropriate form in the target text, thus the application of contrastive lexical and syntactic knowledge. And the **generation** of the target text involves the corresponding problems of style, syntax, and morphology. With the advent of spoken-language translation systems, these can be used to illustrate problems of **speech processing**, both analysis and synthesis. The focus of such systems on task-oriented cooperative dialogues also affords an opportunity to look at issues relating to dialogue and discourse. The multilingual aspect of these issues provides an interesting additional dimension.

#### 4 TEACHING TO TRAINEE TRANSLATORS

The approach of trainee translators to M(A)T is, as opposed to that of CL students, far more practical and considerably less theoretical. Generally, they are interested in MT mainly from the point of view of the user (Balkan, 2001: 8), with a primary focus on the interaction with MAT tools. As stated by Kenny and Way (2001: 15):

Translation students typically have no background in mathematics or statistics, no programming experience, and little or no training in formal or computational linguistics. They can, however, be expected to have excellent command of their source and target languages, and to have the transfer skills required to translate between the two. They are normally well practised researchers, used to getting up to speed in the intricacies of the

specialised areas in which they have to translate. They are also alert to nuance, the importance of cohesion and thematic structure in creating texture, and the roles that textual function, target language audience and text type might play in making low and high-level translation decisions.

The humanistic, non-computational background of translation students entails two main difficulties when it comes to teaching them about MT: first of all, instructors have to fight widely-spread preconceptions and fears on the part of students, who think of machines as job thieves; and second, only a reduced number of students has ever played with an MT system, and many of those who have tend to think that MT output's quality always equals zero, since they do not have a clear idea of what MT systems are really useful for (Yuste Rodrigo, 2001: 45).

To tackle this problem, many instructors suggest starting their MT courses by 'levelling the playing field', «illustrating some of these unfortunate advertising claims, and reporting on and correcting some of the popular misconceptions about MT» (Kenny and Way, 2001: 14). As Somers (forthcoming) puts it: «illustrating how bad translation software can be is a useful precursor to showing the best that it can offer.»

A second step in the teaching of MT to trainee translators would be convincing them of the importance of translation technology in the current market. As stated by Yuste Rodrigo (2001: 45), students should become aware of the fact that (i) «MT systems and applications are essential components of today's global multilingual documentation production»; and (ii) «MT is employed in large multilingual organisations and international companies» which does not necessarily translate into a loss of jobs but rather unexpectedly «opens up new work avenues for translators.» In sum, it is all about proving students that, as Hutchins (1997:1,



apud Yuste Rodrigo, 2001: 47) writes, «computer-based translation systems are not rivals to human translators, but... aids to enable them to increase productivity in technical translation.»

A third step in this process would be, therefore, enabling students hands-on experience with M(A)T applications. According to Gaspari (2001: 42) and Yuste Rodrigo (2001: 48), among others, experience shows that a step-by-step pedagogical approach such as the one described along these lines can change trainees' initial attitude towards a more informed and welcoming standpoint as regard to MT.

In a forthcoming article, Somers distinguishes seven main M(A)T-related tasks of particular interest in the training of future translators, namely: (i) MT evaluation; (ii) assessing MT for assimilation; (iii) post-editing; (iv) drafting guidelines for controlled language; (v) dictionary updating; (vi) simulation of workflow scenarios; and (vii) criticism of documentation and general usability. According to Belam (2001: 32-34) such exercises can help trainees enhance an array of skills, e.g.: (i) text analysis; (ii) SL competence; (iii) translation criticism from a communicative perspective (referred to by him as *Nature of Communication*); (iv) quality evaluation; and (v) linguistic awareness.

## 5 TEACHING TO STUDENTS OF FL

The idea of using MT for FLT purposes is based on the fact that translation is often part of the curriculum for FL learners. Literature

on the topic is still very sparse and mainly based on experimentation with commercially available MT systems<sup>7</sup>. It has been suggested that MT can be used as a medium to highlight onomasiologic differences between L1 and L2 for FL learners; however, many authors remain skeptical. In the middle of this controversy authors like Lewis (1997) regard MT as an interesting tool for enhancing some skills, but put into doubt that it can ever become part of the standard repertoire of FLT.

According to Somers (forthcoming), there are two main applications of MT to FLT. The first one is encouraging students to produce a commented translation, including, on the one hand, a classification of errors from a linguistic or pragmatic perspective, and on the other, some notes on post-editing (if appropriate). The second application is using MT output 'as a bad model', to reinforce student's awareness of onomasiologic differences between L1 and L2. This exercise can be done in two ways - *forwards* (L2 > L1) and *backwards* (L1 > L2). The former is generally considered very useful, since as stated by Anderson (1995) poor quality TTs are often the result of loan translations at lexical and/or syntactic level, which helps students reinforce their linguistic awareness of the two languages. The *backwards* model is more controversial, though. According to many, it can help introduce or reinforce incorrect language habits on the part of the student, as obviously his/her awareness of the L2 is much more reduced than that of the L1. A solution to this problem is provided by Richmond (1994), who suggests using a model translation. According to this author, the implementation of MT into FLT entails, however, further problems: some students may find MT software very difficult to use, which in turn supposes an extra involvement on the part of the instructor.

<sup>7</sup> Cf. Lewis (1997) on the implementation of Power Translator (EN<>GE); Mitkov et al. (1996) on Italian Assistant (EN<>It); Anderson (1995) on Targumatic (HE>EN); Richmond (1994) on French Assistant (EN>FR); and Corness (1985) on ALPS (EN>GE).



Assets and drawbacks aside, MT awareness on the part of FL students is justified, as stated by Lewis (1997: 261, apud Somers, forthcoming) from a real-market perspective, as «future employers may expect prospective graduates in modern languages to have sufficient skills and background knowledge in translation technology to influence decisions on whether or not to invest in MT.»

## 6 CONCLUSIONS

The cross-disciplinary nature of M(A)T poses important challenges on instructors, who have to deal with students from very different backgrounds. According to Somers (2001), three main target groups can be distinguished: students of CL, trainee translators and students of FL. Generally students of CL are concerned with automation from a theoretical and practical point of view, and are familiar with experimental systems; conversely, trainee translators are focused on more practical topics, such as user interaction with commercially available MAT tools, and the impact of technology on their working conditions; the incipient literature on MT applied to FL teaching focuses mainly on the use of MT as a bad model to reinforce the differences between L1 and L2.

Instructors must adapt their teaching strategies to each target group, weighing their assets and weaknesses. As suggested by Kenny and Way (2001: 14), pedagogical approaches should possibly encourage learning by deduction in the case of students of CL, whereas induction should be the starting point for trainee translators and FL students.

Despite the critics of technophobes and so-called *ludists* MT can help students from the three mentioned domains enhance, among others, their computational, analytic, organizational, and language skills. In sum, as Clavier and Poudat (2001: 22) put it, «MT courses help the

language students to master their own translator environment and the linguists to understand their role in computational linguistics.»

## LIST OF ABBREVIATIONS

ACL	Applied Computational Linguistics
AI	Artificial Intelligence
ALPAC	Automatic Language Processing Advisory Committee
ALPS	Automated Language Processing System
CALL	Computer-Aided Language Learning
CL	Computational Linguistics
DCU	Dublin City University
EN	English
FAHQQT	Fully Automatic High Quality Translation
FAT	Fully Automatic Translation
FL	Foreign Languages
FLT	Foreign Language Teaching
FR	French
GE	German
HAMT	Human-Aided Machine Translation
HE	Hebrew
HL	Human Linguistics
IT	Information Technology
IT	Italian
L1	Source Language
L2	Target Language
MAHT	Machine-Aided Human Translation
MAT	Machine Aided Translation
MIT	Massachusetts Institute of Technology
MT	Machine Translation
NLP	Natural Language Processing
ST	Source Text
TS	Translation Studies
TT	Target Text





## ACKNOWLEDGEMENTS

The substantial part of the work presented in this paper was performed within the scope of the following R&D-project: *Diseño de un tipologizador textual para la traducción automática de textos jurídicos (español -inglés, alemán, italiano, árabe)* (PB98-1399, DGICYT, 1999-2002).

The author would like to thank Prof. Harold Somers for inspiration and guidance, and Federico Gaspari and Dimitra Kalantzi for their bibliographic suggestions on the teaching of MT for general purposes and for FL teaching, respectively.

RECIBIDO ENERO DE 2003

## REFERENCES

- Anderson, D. (1995): «Machine Translation as a Tool in Second Language Learning», *CALICO Journal* 13(1), 68-97.
- Balkan, L. (2001): «Exploiting the WWW for MT teaching», in Forcada, M. L., J. A. Pérez-Ortiz, and D. R. Lewis (eds.).
- Belam, J. (2001): «Transferable Skills in an MT Course», in Forcada, M. L., J. A. Pérez-Ortiz, and D. R. Lewis (eds.).
- Clavier, V., and C. Poudat (2001): «Teaching Machine Translation in non Computer Science Subjects: Report of an educational experience within the University of Orleans», in Forcada, M. L., J. A. Pérez-Ortiz, and D. R. Lewis (eds.).
- Corness, P. (1988): MT in the University environment in 1988, In Proceedings of AURA 1988 Conference (pp. 47-61), Brussels.
- EAMT (2001): «What is Machine Translation?», available at: <http://www.eamt.org/mt.html>.
- Forcada, M. L., J. A. Pérez-Ortiz, and D. R. Lewis (eds.) (2001): *MT Summit VIII: Workshop on teaching Machine Translation*. Santiago de Compostela: IAMT, EAMT.
- Gaspari, F. (2001): «Teaching Machine Translation to Trainee Translators: a Survey of Their Knowledge and Opinions», in Forcada, M. L., J. A. Pérez-Ortiz, and D. R. Lewis (eds.).
- Hahn, W. (1995?): «Machine Translation», available at: <http://www.racai.ro/awd/awdr16/hahn.html>
- Hatim, B., I. Mason (1990): *Discourse and the Translator*. London; New York: Longman.
- Hutchins, W.J. and H.L. Somers (1992): *An Introduction to Machine Translation*. London: Academic Press.
- Kenny, D., and A. Way (2001): «Teaching Machine Translation & Translation Technology: A contrastive study», in Forcada, M. L., J. A. Pérez-Ortiz, and D. R. Lewis (eds.).
- Lewis, D. (1997): «Machine Translation in a Modern Languages Curriculum», *Computer Assisted Language Learning* 10, 3, pp. 255-71.
- Mitkov, R., J. Higgins-Cezza, O. Fukutomi (1996): «Towards a More Efficient Use of PC-Based Machine Translation in Education», in *Translating and the Computer 18*. London: Aslib.
- Pérez-Ortiz, J. A., and M. Forcada (2001): «Discovering Machine Translation Strategies Beyond Word-for-Word Translation: a Laboratory Assignment», in Forcada, M. L., J. A. Pérez-Ortiz, and D. R. Lewis (eds.).
- Richmond, I. M. (1994): «Doing it backwards: Using translation software to teach target-language grammaticality», *Computer Assisted Language Learning* 7, 65-78.
- Somers, H. L. (2001): «Three Perspectives on MT in the Classroom», in Forcada, M. L., J. A. Pérez-Ortiz, and D. R. Lewis (eds.).
- Somers, H. L. (forthcoming): «Machine Translation in the Classroom», in H.L. Somers (ed.): *Computers and Translation: A handbook for translators*. Amsterdam; Philadelphia: John Benjamins (Translation Studies).
- Yuste Rodrigo, E. (2001): «Making MT commonplace in Translation Training Curricula – Too Many Misconceptions, So Much Potential!», in Forcada, M. L., J. A. Pérez-Ortiz, and D. R. Lewis (eds.).