Teaching MT - An Indian Perspective

Sivaji Bandyopadhyay
Computer Science & Engineering Department
Jadavpur University
Kolkata, INDIA.
sivaji_ju@vsnl.com

Abstract

The different factors that affect the level of teaching of machine translation, both as a formal courseware and in an informal manner to students doing developmental and research projects can be identified as the background of the student, the state of machine translation research in the country, the language pattern followed in the education system and the motivation for the students in terms of higher studies and the development of the software industry in the multilingual setting. The paper mainly focuses on the teaching of machine translation in an Indian perspective that gives emphasis on teaching the subject to the students of Computer Science and Engineering.

1 Introduction

Machine translation has become a key technology in present-day globalized communications scene, a truly cross-disciplinary technology that should be used with a certain level of understanding. Many universities and academic institutions now teach courses on machine translation, both at graduate and undergraduate level. There are Universities and academic institutions where students carry out doctoral, graduate or undergraduate research and development works in the area of machine translation. These students may not have any formal courseware on machine translation but they are taught the subject by their teachers, albeit informally. Teaching is aimed at trainee translators and students of translation studies, at computational linguistics or computer science students, or even at learners of a foreign language. MT is of interest also to students involved with information technologies in a multilingual setting (electronic commerce, localization, multilingual documentation, cross-language information retrieval, etc.).

The cross-disciplinary nature of machine translation poses important challenges to instructors, whether teaching about linguistics and translation to computer science and engineering students or vice versa, while the existence of widespread preconceptions, fears, and expectations about MT also play a role. On the other hand, with the growth in the internet, commercial MT systems are more readily available than ever, and, if appropriately used, may be very beneficial for education. The world wide web may also be seen as the medium used for teaching, as it may naturally integrate real MT systems as part of the learning environment.

2 Different factors that affect teaching Machine Translation

The different factors that affect the level of teaching of machine translation, both as a formal courseware and in an informal manner to students doing developmental and research projects can be identified as the background of the student, the state of machine translation research in the country, the language pattern followed in the education system, and the motivation for the students in terms of higher studies and the development of the software industry in the multilingual setting.
There is a considerable literature on MT, translation software and Language Engineering in general, but only a small proportion of that literature relates to the teaching of MT and most of it is very recent. Again, that literature is based on the experiences of teaching the subject as a formal courseware. But there are students who usually do various developmental and research projects as part of their course curriculum and they need to be taught the basics of natural language processing in general and machine translation in particular.

**Background of the students**

Students with different backgrounds and interests have different needs with respect to MT teaching. This requires the development of separate courseware and the laboratory work for the different classes of students. The background of the students identifies their expectations from the course, their motivations, the preconceptions they have about the subject, the transferability of skills during an MT course and also what the instructors expect the students to do.

Somers (2001) has identified three different but related perspectives of MT teaching depending on three distinct types of students: students of Computational Linguistics (CL), trainee translators and foreign language learners. Students of Computer Science & Engineering and Information Technology courses pose a different perspective than the students of linguistics. MT offers some interesting computational problems for computer scientists. Translation software offers some good examples in human-computer interaction and interfaces, and specially in software documentation. There is an ever increasing range of multilingual applications which incorporate MT as a component part, and researchers and developers in these areas are also interested in knowing about MT technology.

For the students and teachers of CL, then, MT systems can be used to illustrate problems (and solutions) in language analysis at various levels both monolingually and contrastively. Clavier and Poudat (2001) reports the educational experience in teaching machine translation in non-computer science subjects. In French universities, machine translation is more often taught in Faculties of Arts - Linguistics and Foreign Language Departments - rather than in Faculties of Science - Computer Science Departments.

Translators need some insight into how MT works, why it is difficult, what kinds of translation tasks MT is appropriate for, what alternative computational tools are available and how to integrate them into the workflow. The aim with these students is rather to emphasize the difficulties that the computer faces, often in cases that a human translator takes in their stride. It needs to be taught to these students that MT and CAT systems are best suited to repetitive materials such as manuals whereas literatures can be best left to the translators. Yuste Rodrigo (2001) focuses on the two main difficulties that instructors face while teaching to translation students: first, a misunderstanding of what MT is really useful for, which normally leads to the misconception that MT output's quality always equals zero; second, a widespread fear that machines are to replace human translators, consequently leaving them out of work. In order to fight these generalised prejudices on MT among (future) translators, translation instruction should be primarily practical and realistic, as well as learner-centred.

Reuther (1999) reports the LETRAC survey on translator curricula offered in the countries of the participating project partners, i.e. Denmark, Germany, Greece, Portugal and Spain as well as in other...
European countries. The LETRAC Curriculum Modules meant for translation students is detailed in (Badia et. al., 1999). The curriculum consists of three more or less self-contained modules which are however interrelated in some respect: **Module A** on Introduction to Computer Science, **Module B** on IT and DTP for Translators and **Module C** is on Language Engineering.

MT can be used as Computer Assisted Language Learning (CALL) tool in the teaching of foreign languages. MT software is generally not designed with language learners in mind, so caution must be taken in using it for this purpose. Burnage (1998) looks at the Computer Assisted Language Learning facility for language students that include the use of machine translation techniques.

Kenny and Way (2001) describes the teaching of MT courses to two sets of students with different backgrounds: machine translation course to undergraduate students in Applied Computational Linguistics and Computer-Assisted Translation (CAT) course in two translator-training programs, one undergraduate and another postgraduate. The students are made aware that all systems can be expected to show an overall improvement in quality if notions of sublanguage and controlled language are taken into account. Students of Computational linguistics are viewed primarily as developers of MT systems. Translation students are more interested in CAT, and specially TM tools.

Belam (2001) addresses the transferability of skills during an MT course mainly designed for students of modern languages. Transferable skills refer to the acquisition of abilities that are required for the mastery of the subject to be taught but also have a wider application in other areas. The transferable skills that a student of Computer Science and Engineering or Information Technology expects to learn from an MT course are the language analysis techniques, the database and the knowledge representation techniques for non-numeric computation. The students learn the skill of linguistic awareness required when 'localising', as opposed to translating, documentation produced in one country or culture for use in another.

### State of Research in MT

The second important factor is the state of research in machine translation in the languages of that country. This is an important motivating factor for teachers to start teaching the subject formally or informally as well as for students to start learning the subject. The background of the researchers determines the class of students to whom MT will be taught by them in the initial phases. Later on, as the research on MT becomes truly multidisciplinary in nature, it will be taught to a wide variety of students. The level of research in turn determines whether on-line research MT systems in the native languages are available or not which can serve as a teaching aid. Availability of commercial MT systems also depends on the level of research in the country. Continuous academic and research interactions with various Universities and Institutions worldwide increases the level of research work in the country. It also works as a motivating factor for students since it increases the possibility of their future higher studies.

The most important tools as far as MT teaching is concerned are the MT/MAT tools themselves. MAT tools include the dictionaries, terminologies, thesauri, bilingual concordances and translation memory systems. Blanc (2001) describes an interactive hypertextual environment for MT training. The main tool used for MT research at GETA (Groupe d'Etude pour la Traduction Automatique), France is the ARIANE-G5 MT shell, a generator of MT.
systems. The ARIANE-G5 shell and the methodology developed at GETA have been utilized in the 15-language Universal Networking Language project initiated by the Institute of Advanced Studies of the United Nations University (http://www.ias.unu.edu). Hindi, the national language of India, is included in the UNL project and the Indian Institute of Technology, Bombay is the participating institute. The interactive hypertextual environment for the control of ARIANE-G5, CASH (for Control of Ariane Supported by HyperCard) was initially developed as a tool for the MT developer and proved very useful for the training of beginners in the MT field.

Forcada (2000) has proposed a laboratory assignment for non-computer science majors in a translation course using commercial machine translation systems to discover the word reordering patterns used by the commercial MT systems. The assignment is suitable for computer science majors also. A laboratory assignment has been proposed in (Perez-Ortiz and Forcada, 2001) to study the way in which some commercial machine translation programs translate whole sentences and how the translation differs from a word-for-word translation.

Balkan (2001) has given an overview of what the World Wide Web currently offers in the way of resources for teachers and learners of machine translation and discusses how the potential can be further exploited in the future.

**Language Pattern in the Education System**

The third important factor that guides teaching machine translation is the language pattern that students usually follow in the education system of the country. Let us take the example for India. The students are generally taught in their mother tongue up to the school level but they also learn English as their second language. They generally translate English sentences to their mother tongue to learn the meaning and also do a translation while writing a sentence in English. The translation seems similar to an example based machine translation technique. There are students who learn in English medium schools but they are also exposed to their mother tongue while in their family or in other gatherings. The exposure of two languages since school days prepares the students at least as amateur translators. Moreover, the students usually learn Hindi, the national language of India, through television, films and other media. Languages are taught to the school students with strong grammar content. The students are thus able to understand the intricacies of Rule based machine translation systems. Higher education in India is taught in English medium.

The students need not learn a foreign language during the course on MT if they know at least two languages since their school days and there are machine translation systems, either commercial or research, available in those languages.

**Motivation for the Students**

The fourth factor is the motivation for the students in terms of higher studies and the development of the software industry in the multilingual setting where the students can possibly be absorbed after their course. This potential must exist not only within the country but also in the global frame. This necessitates that the students must learn a foreign language beside English. This can be made optional during the course and the students can learn a language afterwards.

Abaitua (1999) examines why it is worth learning translation technology. Software localization is the paradigm of the need for translation technology. The main role of localization companies is to help software
publishers, hardware manufacturers and telecommunications companies with versions of their software, documentation, marketing and web-based information in different languages for simultaneous worldwide release. The Localization Industry Standards Association (LISA) Education Initiative Taskforce (LEIT) is a consortium of schools training translators and computational linguistics that was announced in 1998 as an initiative to develop a promotional program for the academic communities in Europe, North America and Asia. The National Centre for Software Technology (NCST), Mumbai, India is one of the academic institutions involved in the first release of LEIT. The main goal of the LEIT initiative is to introduce localization courseware into translation studies.

Technology has been playing an increasingly important role within language translation over the past five decades or so, and at present its impact is undisputedly extensive and has reached an unprecedented climax that deserves careful consideration as a crucial factor that affects human translators in the first place. The Translation technology industry is now moving in the direction of integrating systems that combines term management, TM, MT and related tools to create a seamless full service localization environment. Translators need to become acquainted with technology, because good use of technology will make their jobs more competitive and satisfactory. The limits to technology arise when going beyond the mechanical and routine aspects of language.

The translation profession in general, and particularly so in many countries like Italy (Gaspari, 2001) does not seem to be taking the situation adequately into account, so as to lag behind the ever-accelerating pace of technological development in the present day multilingual information society. Gaspari (2001) reports upon a survey carried out among some trainee translators who took courses on machine translation. The students in general showed a more balanced attitude and a shift towards a more positive perception of MT after their course was over.

3 Teaching Machine Translation - An Indian Perspective

India is a vast multilingual country where there are 18 scheduled languages besides several others in the different parts of the country in which students learn at least up to the primary level. Computerization in India has mainly concentrated on using English as the interfacing language. Around 1990, the Government of India has started the Technology Development in Indian Languages (TDIL) project after which the work on computerization of Indian languages got a big fillip. Currently, language resources in several Indian languages are being developed and some of these resources are meant for machine translation.

In a large multi-lingual society like India, there is a great demand for translation of documents from one language to another. Most of the state governments work in the respective regional languages whereas the Union Government's official documents and reports are in bilingual form (Hindi/English). In order to have a proper communication there is a need to translate these reports and documents in the respective regional languages. With the limitations of human translators most of this information (reports and documents) is missing and not percolating down. A machine assisted translation system or a translator's workstation would increase the efficiency of the human translators.

This paper focuses on the Indian perspective of teaching machine translation to the students. The history of machine translation in India is not very old and dates back to 1984 when research work on
machine translation in India languages started at the Indian Institute of Technology, Kanpur. Since then several academic institutions in India has started research work in machine translation in Indian languages, Indian languages to English and English to Indian languages. The subject is taught as a formal courseware in some Universities and Institutions in the Computer Science and Engineering department or in the course on computational linguistics whereas there are many Universities and Institutions that carry out projects on machine translation for the graduate and undergraduate students of Computer Science and Engineering as part of their course curriculum. Doctoral students in the same field also carry out research in the area of machine translation though they may not have any formal training on the subject. The students learn the subject informally from their guides while doing their projects.

Since most of the Universities and institutions teach machine translation to students of Computer Science & Engineering, the course content is designed that suits such students. India has a good number of high quality linguists and computer engineers but the concept of computational linguistics as a multidisciplinary field of study is something new. There are very few institutions that teach machine translation to the students of computational linguistics. The students share the common misperceptions about machine translation as they are not generally exposed to any commercial or research machine translation system. Currently, two research machine translation systems are available on-line in the English-Hindi domain and research work are also going on in English to other Indian languages and vice versa. The necessity that the students must learn a foreign language is now not so stringent. But the knowledge of a foreign language besides English will obviously prepare the students better for the global multilingual software industry.

Research on machine translation is mainly carried out by scholars and teachers with a Computer Science & Engineering background. The research has now reached a state when students are feeling interested in taking up developmental or research projects in the area. The teachers have started teaching the subject to the students mainly with Computer Science and Engineering background either as a formal courseware or informally. Though there are no commercial MT systems in Indian languages or with English and Indian languages there are two on-line research MT systems that can be used for teaching. The most important factor is that the Universities or Institutions where teaching and/or research in machine translation are going on are in active collaboration with the most prestigious Universities or Institutions in the world. This has increased the possibilities of higher studies in the relevant field for the students. Otherwise, the scope for higher studies and the scope for good jobs in high profile companies are not so much for the computer engineers who study machine translation. It is a very difficult task to motivate the students to pursue research career in the area of machine translation.

A brief account of the state of machine translation research is outlined below. A more detailed report can be found in (Bandyopadhyay, 2000) and in the relevant web sites of the various institutes as mentioned there.

The research on machine translation in India is mainly funded by the Government agencies. The Ministry of Information Technology, Government of India, (http://www.mit.gov.in/) has identified the following domains for development of domain specific translation systems: government administrative procedures and formats, parliamentary questions and
answers, pharmaceutical information, legal terminology and important judgments, and so on. The Ministry initiated the TDIL (Technology Development for Indian Languages) project in 1990-91 to support R&D efforts in the area of Information processing in Indian languages covering machine translation among others.

A machine aided translation system (Anusaaraka) among Indian languages has been built with funding from TDIL project. The Anusaaraka system presents an image of the source text in a language close to the target language. Anusaarakas has been built for five pairs of languages: Telugu, Kannada, Marathi, Bengali and Punjabi to Hindi. They are available for use through e-mail servers. The web page where relevant details are stored is maintained at http://www.iiit.net/ltrc/Anusaaraka/anu_mail_help.html. Work is going on in building an English to Hindi Anusaaraka system, which will be a test of building a system between two languages that are far apart. The system so developed will be available as free open-source software under GPL. The work on the Anusaaraka project started at the Indian Institute of Technology, Kanpur. It is now being carried out at the Language Technologies Research Center, Indian Institute of Information Technology, Hyderabad (http://www.iiit.net/research/ltrc).

The Natural Language group of the Knowledge Based Computer Systems (KBCS) division at the National Centre for Software Technology (NCST), Mumbai is working on MaTra, a human-aided transfer-based translation system for English to Hindi. The work is supported under the TDIL project. The domain being explored is news, but the approach is applicable to any domain. The system breaks an English sentence into chunks, analyzes the structure and displays it using an intuitive browser-like representation, which the user can verify and correct, after which the system generates the Hindi equivalent. A prototype that can translate simple (single verb group) sentences has been developed. The group is currently developing a practical framework for the syntactic transfer of compound-complex sentences from English to Hindi. NCST maintains an on-line machine translation facility in the Matra system (http://konark.ncst.ernet.in/~matra). It accepts input English sentences from the user and provides the necessary output in Hindi in a downloadable font. More information about the system can be obtained from the NCST Website (http://www.ncst.ernet.in/kbcs/NLP.html).

The Machine Aided Translation System ANGLABHARATI, for translation from English to Hindi for the specific domain of Public Health Campaign has been developed. This technology is proposed to be extended to another domain for translation of Financial and Supplementary Rules of Government of India and related correspondence. The ANGLABHARATI project was launched at the Indian Institute of Technology, Kanpur in 1991 for machine aided translation from English to Indian languages.

The Multilingual Pocket Translator Design Project was undertaken by the Center for Development of Advanced Computing (CDAC) with a view to foreign travelers visiting India. The same pocket translator is useful when a person moves from one state to another within India. Further details about the Pocket Translator can be obtained from the website http://vishwabharat.tdil.gov.in/pocket_trans.htm and also from the CDAC website http://www.cdac.org.in/. Work is also going on at the CDAC for the development of an English-Hindi machine translation system using Tree Adjoining Grammar.

Work on a knowledge-driven generalized example-based Machine Translation system from English to Indian languages is being carried out in the ANUBAAD.
Project at the Computer Science and Engineering Department, Jadavpur University, Kolkata. It is currently translating short single paragraph news items from English to Bengali using a hybrid approach. Headlines are translated using knowledge bases and example structures. The sentences in the news body are translated by the application of hybrid techniques. The methodologies can be used for developing similar systems for other Indian languages.

4 Teaching Machine Translation - The Indian Scenario

There are a number of premiere institutions in India which teach machine translation either as part of their curriculum on Computer Science and Engineering or on Computational Linguistics. All the institutions teach Artificial Intelligence as part of their curriculum on Computer Science & Engineering. The Indian Institute of Technology, Kanpur (http://www.cse.iitk.ac.in) provides courses on Introduction to Natural Language Processing, NLP Semantics and Machine Translation as part of their Post graduate programme in Computer Science & Engineering. The following topics are covered in the course on Machine Translation: Overview of Natural Language Processing; Syntax, semantics, context and world of knowledge; Strategies for machine translation, Direct, Transfer and Interlingua approaches; Rule based, Example based and Hybrid Methodologies; Construction of lexical data-base, Text generation, Machine-aided translation, user interfaces; Examples of English-Hindi and Hindi-English machine translation.

The Indian Institute of Technology, Bombay (http://www.cse.iitb.ac.in) offers courses on Natural Language Processing and Cognitive Psychology as part of their undergraduate programme on Computer Science & Engineering. The course on NLP includes Machine Translation, Natural Language Interfaces, Natural Language Generation along with other relevant topics. The National Centre for Software Technology, Mumbai (http://www.ncst.emet.in) teaches MT as part of the Natural Language Technology module in its Advanced Post Graduate Diploma for Software Technology course.

The Language Technologies Research Centre (LTRC) in the International Institute of Information Technology, Hyderabad (http://www.iiit.net) conducts several research programmes and participates in the undergraduate and postgraduate programmes at the IIIT. Courses for specialization in Artificial Intelligence / Natural Language Processing (AI/NLP) stream for four year B.Tech Honours programme and two year M.Tech programme are conducted by LTRC. The following research programmes are offered by the LTRC : MS by Research in Computer Science (with specialization in NLP), Ph.D. in Computer Science (with specialization in NLP), One Year Post Graduate Diploma in Computational Linguistics and Ph.D. in Computational Linguistics. Incidentally, this is the only institute in India that offers programmes in Computational Linguistics. LTRC offers several courses in this area including Natural Language Processing Introduction, Natural Language Semantics, NLP Applications: Machine Translation and Information Extraction and Statistical Methods in NLP.

Experience at Jadavpur University

The subject of machine translation is not yet introduced as a course of study in the Computer Science & Engineering department at the Jadavpur University. There is a research group on MT which includes doctoral, post-graduate and undergraduate students. The work has been initiated by the author in the year 1997. The students who work in the group have
interesting exposure to manual translation based on example based techniques. They studied in their mother tongue up to the school level but English was their second language. They were taught English during their school days by translation method into their mother tongue. At the same time they were good in English grammar. While preparing an answer in English, they first try to frame it in their mother tongue and then translate that to English. This practice somehow receded when they joined the Computer Science and Engineering department. But they continued to follow the practice of manual translation while reading from English texts or comprehending other English documents. It has been observed that the students basically follow a hybrid approach to translation; following example based techniques in most of the times and using their knowledge of grammar when necessary. In fact, these students understand clearly the Example based approach of translation than other approaches and we have been following the Hybrid approach of machine translation in our translation works from English to Bengali, an important language in India.

On the other hand, these students study Artificial Intelligence as a course of study where the main focus is on search techniques, knowledge representation techniques and the resolution principle. The section on natural language processing is covered during their AI course only sketchily. These things add to the problem of teaching machine translation. Since there is no formal courseware on machine translation the students are taught from several survey articles, text books, research papers, conference proceedings and the web. The students are mastering the subject to an appreciable extent. Since there is a scarcity of translation resources and tools most of the works of undergraduate and postgraduate students are oriented towards designing and developing the resources and the tools.

The MT research group also conducts research in the area of Natural Language Interfaces, Question Answering Systems and Automatic Summarization among others that have strong applications of example based machine translation techniques. The Postgraduate programme in the Computer Science & Engineering department has recently been revised and it now includes courses in Natural Language Processing and Computer Processing of Indian Languages. Machine Translation is to be taught in both these courses. The revised programme is going to be introduced from the next academic session.

5 Conclusion

The time has come when India should have degree programs in computational linguistics and machine translation should be taught as a regular subject in that. The government and the Indian industry should come forward to expand the computerization in Indian languages. The common people in India do not know English and it is high time that the effects of computerization should reach them. More collaboration is necessary with eminent educational institutions all over the world so that Indian computational linguists are a part of the global fraternity.

References

1. Abaitua, Joseba (1999) "Is it worth learning translation technology?" 3rd Forum on Translation in Vic. Training Translators and Interpreters: New Directions for the Millennium, Universitat de Vic, Spain, http://sirio.deusto.es/abaitua/konzeptu/ta/vic.htm.

2. Badia, Toni, et. al. (1999) "LETRAC Curriculum Modules", Deliverable D3, European Commission DG XIII Telematics Application Programme Fourth
3. Balkan, Lorna (2001) "Exploiting the WWW for MT teaching", MT SUMMIT VIII Workshop on Teaching Machine Translation, Santiago de Compostela, pages 7-11.

4. Bandyopadhyay, Sivaji (2000) "State and Role of Machine Translation in India", Machine Translation Review, Issue No. 11, December 2000, pages 25-27.

5. Belam, Judith (2001) "Transferable skills in an MT Course", MT SUMMIT VIII Workshop on Teaching Machine Translation, Santiago de Compostela, pages 31-34.

6. Blanc, Etienne (2001) "An Interactive Hypertextual Environment for MT Training", MT SUMMIT VIII Workshop on Teaching Machine Translation, Santiago de Compostela, pages 51-55.

7. Burnage, Gavin (1998) "Teachers and Technicians: working together for effective use of Information Technology in Language and Literature", in Sarah Porter and Stuart Sutherland (eds) Teaching European Literature and Culture with Communication and Information Technologies: Selected Papers, Oxford: CTI, Oxford University Computing Library, http://info.ox.ac.uk/ctitext/publish/occas/eurlit/burnage.html.

8. Clavier, Viviane and Celine Poudat (2001) "Teaching Machine Translation in non Computer Science Subjects: Reports of An Educational Experience within the University of Orleans", MT SUMMIT VIII Workshop on Teaching Machine Translation, Santiago de Compostela, pages 19-23.

9. Forcada, Mikel L. (2000) "Learning machine translation strategies using commercial systems: discovering word-reordering rules", MT 2000: Machine Translation and Multilingual Applications in the New Millennium, Exeter, UK, pages 7-18.

10. Gaspari, Federico (2001) "Teaching Machine Translation to Trainee Translators: a Survey of Their Knowledge and Opinions", MT SUMMIT VIII Workshop on Teaching Machine Translation, Santiago de Compostela, pages 35-44.

11. Kenny, Dorothy and Andy Way (2001) "Teaching Machine Translation & Translation Technology: A Contrastive Study", MT SUMMIT VIII Workshop on Teaching Machine Translation, Santiago de Compostela, pages 13-17.

12. Perez-Ortiz, Juan Antonio and Mikel L. Forcada (2001) "Discovering Machine Translation Strategies Beyond Word-for-Word Translation: a Laboratory Assignment", MT SUMMIT VIII Workshop on Teaching Machine Translation, Santiago de Compostela, pages 57-60.

13. Reuther, Ursula (ed.) (1999) "LETRAC survey findings in the Educational Context", Deliverable D1.2, European Commission DG XIII Telematics Application Programme Fourth Framework, project LE4-8324, www.iai.uni-sb.de/letrac/D12.doc.

14. Somers, Harold (2001) "Three Perspectives on MT in the Classroom", MT SUMMIT VIII Workshop on Teaching Machine Translation, Santiago de Compostela, pages 25-29.

15. Yuste Rodrigo, Elia (2001) "Making MT Commonplace in Translation Training Curricula-Too Many Misconceptions, So Much Potential!", MT SUMMIT VIII Workshop on Teaching Machine Translation, Santiago de Compostela, pages 45-49.