Democratizing translation technologies
The role of humanistic research

Anthony Pym

Intercultural Studies Group, Rovira i Virgili University
Avda Catalunya, 35. 43002 Tarragona, Spain
`anthony.pym@urv.cat`

> "the gap between the underlying population and those who design the blueprints for rationalization, who lay out production, who make the inventions and discoveries which accelerate technological progress, becomes daily more conspicuous" (Marcuse [1])

Abstract— **Recent research on translation memories and machine translation technologies tends to focus on technical issues only, falsely abstracting the technologies from the many different social situations in which they are ostensibly to be used. At the same time, the revolutionary promise of the systems with learning potential is that they will improve output only with widespread use, and thus only through the involvement of different groups of social users. In principle, humanistic research is well positioned to investigate and communicate between the various users, with awareness of different kinds of user, collaborative workflows, text types, and translation purposes. If knowledge on those variables can be fed back into the technical research and development, humanistic research could thus play a key role in enhancing not only the social impact of the technologies, but also their democratization.**

Keywords— **Translation memories, machine translation, sociology of technology, Translation Studies, research methodology**

I. INTRODUCTION

Translation is increasingly carried using translation memory systems (TM) that incorporate machine translation (MT), thus giving TM/MT systems. The MT component in most contemporary systems is moreover statistical or data-based, in addition to various more narrowly linguistic algorithms. For the perspective adopted in this paper, the most important features of the resulting TM/MT systems are the following:

1. In principle, the more you use them, the better they get. This is what I am calling the "learning" dimension of TM/MT.
2. In principle, the more they are online ("in the cloud" or on data bases external to the user), the more they become accessible to a wider range of public users.

These two features are clearly related in that the greater the accessibility, the greater the potential use, and the greater the likelihood of improved performance. In short, these features support each other and could constitute something like a revolution, not so much in the translation technologies themselves, but in the social use and function of translation.

This paper will explore some of the possibilities for research on this sea-change. My basic argument is that the revolutionary potential of the technologies can only be realized if they connect with significantly developed *human* skill sets. That is, if accessibility is to become use, which becomes productivity and quality, *a very wide range of people* are going to have to help connect those dots. This is because, in order to render usable translations, these systems still require either pre-editing of texts to be translated or post-editing after completion of the translation. That is, they require human skills that are not only mostly ignored in the technical literature, but have been the object of very little empirical research and have yet to connect seriously with the training of language professionals.

If I am right, there is a lot to be done, and the people who have to do it are not all technicians.

II. WHAT WE KNOW SO FAR

Let us assume that there are basically three research communities involved here. The first would be the people actually developing technologies in private companies. The second group would comprise the engineers, linguists, and increasingly mathematicians who do research within university environments, with some crossovers and joint projects with industry. The third group, often forgotten or simply not taken seriously, would then be university researchers more firmly within the humanities, usually with backgrounds in linguistics, communication studies, or translation. This third group is my main concern here. It is special because, first, its main contribution to the other groups is supposedly through the provision of training: they are the ones closest to the institutions that are paid to train the translators of the future. When the training is irrelevant, then so might be the humanistic academics. But this group is also key because, as I

shall argue here, they are the ones best positioned to investigate the human aspects of translation technology, including precisely the features that might make the technology revolutionary.

Within the first two communities, research on translation technologies has generally been on technical, linguistic, and quality issues, almost always overlooking the human and training aspects. As Andy Way [2] argues, the developers of data-based MT, whether in industry or the academy, typically "do not seek to address any community other than their own, for they do not feel any need to do so." Way is primarily talking here about the lack of communication within the first two communities, particularly between different groups of TM/ MT researchers. Needless to say, there has been even less concern with addressing the third research community, the "soft" linguists and trainers, who have themselves long abandoned any hope of understanding what is happening in the MT systems as such.

There is nevertheless a sizeable body of research that has been conducted on the human dimension of TM/MT. This has tended to be carried out by humanistic academics, basically as interested followers of the technical developments, mostly without concern for new training needs. Here I focus in the research that has been carried out by this loose third group, specifically with respect to 1) the use of translation memories, 2) the use of machine translation with translation memories, and 3) workflows that can be considered non-professional or "collaborative".

My quick survey here is neither exhaustive nor carefully representative. Yet it might be indicative of some general trends, a few general problems, and perhaps even a handful of pointers for the way forward.

A. Research on the use of TMs

The first studies on the actual use of translation memories tended to be abstract lists of advantages and disadvantages ([3], [4]). Reports of productivity gains were made but varied widely, ostensibly because of the many different variables involved: up to 70% [5], 30% on average [6], or between 2% and 16% [7] (cf. Yamada [8]). Numerous doubts have since been raised about the quality of outgoing translations ([9], [10], [11]), about the cognitive effects of forced segmentation [7], about translation memories propagating errors ([12], [13]), and more generally about communication that occurs without reference to the communication participants ([14], [15], [16], [17]). Some of this discourse might be seen as resistance on the part of those academics who have a deep long-term allegiance to the previous technologies of book culture, the printing press, and the competent human translator. Such a characterization would not seem altogether fair, however, since there are many other kinds of projects underway. For example, Pym ([17], [18]) has argued that cheap, low-quality translations can play an ethically laudable social role. Or again, Yamada [8] reports that productivity gains are significantly higher when a TM has literal translations, which may suggest that the more sophisticated proce-

dures employed by highly trained translators become relatively unproductive in this environment. Not all academics defend the standards they were trained in.

Whatever the underlying ideologies, all these studies compare TMs with fully human translation, either explicitly or implicitly. They voice doubts that circulate among the various communities of users: beyond the company hype, is there real evidence of productivity gains?; what has happened to quality in human terms?; who can tell us when and where TMs are truly beneficial?

The only really negative part of the story is that, after some decades of attempts, we have no firm answers to any of these questions.

B. Research on the use of TM/MT

An early evaluation of TM/MT with post-editing is Guerra Martínez [19]. A comparative evaluation of TM and MT is reported by Guerberof [20] [21], suggesting that MT may be more advantageous in certain situations. More recent evaluations include the "training" of MT databases by feeding in specialized terms and phraseology. From the edge of industry, Thicke [22] reports that "untrained MT is two times faster to post-edit than to translate from scratch; trained MT is three times faster; and trained MT with source control is four times faster". That is, for a handful of subjects, working on a particular text type, under certain conditions, and for a certain kind of quality expectation. One quickly cites the neat general findings and forgets about the many specific variables that produced them.

The method of evaluating translation technologies has been an issue for many years ([23], [24], [25], [26], [27]). There is general agreement that increased productivity is only part of the story. The BLEU algorithm [28] has been used with some success but has also been criticized for excluding the human dimension [29]. Post-editing time is preferred as an industry-relevant yardstick in Thicke [22]. The general issue of quality is problematic in most areas of Translation Studies, and remains so here. I return to this below.

Although the theoretical implications of TM for the training of translators have been mentioned for some time ([14], [16], [17], [30], [31], [32], [33], [34]), it seems that only O'Brien [35] has looked at the specific pedagogical demands of post-editing, while some related but indirect insight can be drawn from Mossop's [36] work on revising and editing in general. TAUS [37] presents results from a workshop on best practices in post-editing, which may be adapted to teaching purposes. Empirical research in teaching environments, with very short-term training in TM/MT, has found that initial use of TM/MT presents slight advantages over fully human translation ([11], [38]). Strangely, I am aware of no research on the teaching of MT use.

C. Research on volunteer and collaborative translation

The use of TM/MT by non-professionals seems to be such a recent phenomenon that there is no fixed name for the

practice. There has been considerable talk about "community translation", "crowdsourcing", "collaborative translation", "untrained translation", "paraprofessional translation", user-generated translation, or, when someone just gave up and pre-ferred tech-talk, "TC3" (standing for "community, crowdsourced and collaborative translation") (cf. comments on all these terms in Pym [39]). Here I opt for the variant "volunteer translation", since I assume that the fundamental difference at stake is the monetary payment received (or not received) by the translator. If a professional translator is one who receives monetary reward, then the opposite term should be "volunteer" (qualifying the person, not the action). Many of the alternative terms here seem shot through with activist ideologies, all of which are very well meant, and none of which highlight the most problematic feature.

If the work of unpaid translators can be labeled this way, then the term "collaborative translation" can refer to workflows where volunteers intervene alongside professionals, presumably in different phases of the workflow.

Understood as such, collaborative translation can be traced back to Holz-Mänttäri [40], who saw translators as "experts in intercultural communication" working to help experts in the various technical fields concerned. In 2001 an interesting pedagogical experiment was carried out between the faculties of Law and Translation at the University of Granada, where students from both faculties worked together on translation projects [41], thus creating a simulacrum of possible collaborative scenarios. The more recent technologi-cal aspects of collaborative translation workflows are reported on by Abekawa et al. [42], Lin et al. [43] and Carson-Berndsen et al. [44], as well as numerous media reports of success in cases like Facebook. There is nevertheless little hard quantitative data on measures of productivity and quality. O'Hagan [45] interestingly sees collaborative workflows as a new training environment. The potential impact on the translation profession is assessed in García ([46], [47]) and Pym [48], but again without extensive empirical data.

In sum, the three fields of interest here have been the object of increasing research in recent years, but only a few papers pay full attention to the human aspects, and next to nothing has been done on the new pedagogical demands.

III. PROBLEMS FOR THE CREATION OF KNOWLEDGE

As might be apparent from this quick overview, humanistic researchers are neither indifferent nor entirely ignorant of the advances in translation technologies, and some of them have been asking interesting questions for quite some time. There are nevertheless a few fundamental problems involved whenever academics try to grapple with concerns that come from industry. Before trying to envisage a way for-ward, consider some of the difficulties we have to live with: institutional belatedness, limits in our scientific training, and fundamentally different approaches to quality.

A. Institutional belatedness

In 2010 Sánchez-Gijón et al. [49] and Torres-Hostench et al. [50] published preliminary reports on a large-scale empirical research project on TMs. Given that this was a four-year research project for which the funding application was probably written one year previously, we might assume that the conceptualization of the research occurred some five years prior to the published reports. That sort of delay is quite normal in academia, at least in the humanities. At the same time, five years can be an eternity in TM/MT development: by the time the researchers have discovered how student transla-tors actually interact with Trados (in the case of this particular project), that particular version of Trados has become out-dated and real-world translators are negotiating integrated MT, hopefully with cheaper and more agile tools. The data is thus condemned to be outdated as soon as it is published.

This relatively long problem-solution cycle should partly explain why a good many academic researchers tend to focus on problems that are quite general and distanced from immediate application. When Yamada [8] reports on the effects of using two different kinds of translation pairs, his question is as eternal as the binarisms of classical translation theory, going back to Cicero's *ut orator* vs. *ut interpres*. The field of research is TM, but the problems and solutions address a millennial debate.

In a sense, academic institutions are academic precisely in order to reach that order of generality. Our institu-tions condemn us to it; we are not really able, and very rarely financed, to produce the technologies of tomorrow, but we might be able to guess at how technologies develop in terms of decades, and how they can affect communication. Of course, such questions are usually too vague and too vast to interest industry, which is precisely why we are not in industry.

B. We don't understand mathematics

Academics in the humanities might have a fighting chance when it comes to understanding linguistic algorithms; we should grasp the basics of probabilistic statistics; and we perhaps have fading memories of calculus and things like that. But we are not getting far with the kinds of statistics that have allowed data-based MT to advance. And it is time we admit-ted it.

As noted above, Way [2] has broadly blamed the statistical MT community itself for not adequately explaining what it does. That may be the case. However, there is a disarming paradox within Way's text itself. Perhaps out of despair with the ignorance of his peers, Way offers a succinct and very well-intentioned explanation of the basic models involved, using quite accessible mathematical notation. Unfortunately the Translation Studies journal in which his text was published has six mathematical symbols missing from that explanation. A copyeditor or editor was presumably so unaware of what was happening that the errors were not picked up and all we see are those blank white squares of non-

codification. Better to admit that we (as writers or copyeditors) do not know what we don't know, to let the developers do their developing, and to stick to the questions that are more properly ours. To which I return below.

C. Epistemologies of quality

As seen in the overview above, the question of quality is one of the main points at which many different roads diverge. For the people developing technologies, a mechanistic same-for-all test like the BLEU scores is objective and meaningful – they want to improve their performance scores on neutral metrics, and they are certainly right to do so. Quality, on this view, is a matter of comparing one set of products with another, and doing so with respect to human-less performance (like when music is performed without an audience – to check that the notes are in the right order).

However, for the people providing language services, and indeed for any organization these days, quality is more clearly a question of having enough testing and control procedures in place. Quality ensues from *regulated* processes, on the assumption that if the process is right, the products are bound to meet standards of quality [51]. If an algorithm produces something of inferior quality, then the solution is to identify the inferiority and root it out. Quality is reduced to quality control, and can be enhanced by multiplying the measures of control. True, this could formally resemble the processes by which data-based MT generates multiple possible translations then selects the most likely one, but there remains a difference between the application of algorithms and the essentially human quality control operative in companies and industry norms. In the company, what matters most is the amount of *time* invested, be it in learning curves, pre-editing, postediting, or making excuses to the client. In the technical metrics, time is not usually a variable.

On the other hand, for anyone in the humanities, quality is more likely to be a question of relative prestige (what *discursive* standards a text meets) and, more interestingly, what people want to do with a text (basic usability). Since the 1970s, translation theories have incorporated text types as a significant variable: different texts require different translation methods, to meet different kinds of discursive norms or ways of representing cultural otherness. Since the mid 1980s, translation theory has generally also recognized that the one text can be translated in different ways to suit different end purposes, thus making the client and the actual use of the translation into significant variables affecting the translation process. This aspect is something that companies are intuitively aware of but only rarely incorporate into quality-control processes. It is something that mechanical quality metrics rigorously exclude. Hence the existence of huge areas where humanistic research might yet find something to say.

IV. CONCLUSIONS

My ultimate concern here is to map out the kind of research that academics in the humanities might best be able to tackle. As should be clear from the above, my position is not that the humanities should simply stop being human, forcing us all to do mathematical linguistics or build operative machines. I want to do research, but I want to keep it within the fields in which we might have some knowledge and relative expertise. Here are some suggestions.

If we are involved in experimental research, then the variables we might want to look at and combine are:

1. Subject variables (technical experience, language training, translator training, awareness of provenance of matches)
2. Work flows (pre-editing, multiple post-editing, collaboration between professionals and volunteers)
3. Time pressure
4. Text types
5. Translation instructions
6. TM/MT programs (different architectures, interfaces, types of memories, types of translations in memories)
7. Languages and directionalities.

I have put these in the order of what could be seen as decreasing importance. Subject variables are first because the possible revolution will depend on volunteer translators and others with minimal specialized training. Of course, this is also the variable with respect to which we might test the effects of specialized training. The effects and modes of pre-editing and post-editing are then of related importance, since there are many possible workflows involved, and many different situations in which one or the other may be more recommendable. The remaining variables are then perhaps of more specific interest: time pressure might connect with the different workflows, since the basic variable in question is time-on-task and its relation to productivity. Text types and client instructions are then related to issues of quality, but there are so many possible text types and diverse translation functions that the findings here should respond to quite specific problems. The comparing of different TM/MT programs is then in penultimate place because, first, this is an industrial and commercial area subject to more hype than knowledge, and second, the programs change so quickly that their relative advantages and disadvantages are likely to be transitory. Language and directionalities are last because I assume that the statistical principles that work in one language pair will eventually work in all others, if and when the databases can be built up. At the same time, anything that can be done to test that naïve assumption, and to help build the databases by generating public or institutional interest, is more than welcome.

The variables have thus been ordered in a rough progression from the most general to the most specific, with my clear preference being for the variables at the more general pole. Note, though, that what is "general" here is also the pole that is most subjective, most oriented toward the peculiarities of people as social actors. This, again, is why we are in the

humanities: our generality should strive to remain human, and draw out the diversity of that condition.

It should be clear that any one of these variables can be combined with any other, or set of others. For example, we might choose to study novice vs. professional with respect to time pressure, or text types, or different instructions. Using this logic, there are numerous possible combinations and the existing research has only scratched the surface. Young researchers should never believe "it's all been done". For that matter, the existing research is on such small groups of subjects (I reckon that all the experiment groups together total less than 400 subjects) that almost all the findings warrant checking by replication.

The measuring of many of these variables can these days be carried out either by traditional linguistics or, more interestingly, by the tools used in process research: think-aloud protocols, keystroke logging, screen recording, and eye-tracking (with stress measured by pupil dilation).

The one variable I have not mentioned here is quality. In a sense, criteria for quality are written into the "different translation instructions" and perhaps into "text type" (since there is a loose correlation between text type and expected function). However, the quality variable is more generally what is to be discovered. That is, we change the text type, translation experience or whatever, and we observe the subsequent effects on the quality of the translation output. Exactly how one measures quality then depends on the aims of the experiment. My personal preference is for a model of "effability" that assumes *everything can always be translated adequately for the purpose, if and when we have enough time to work on it* (cf. Katz [52]). You then measure the time necessary to pre-edit or perform multiple post-edits until adequate quality is deemed to have been obtained (as in Thicke [22]). This does not solve all the problems (post-editors are still subjective), but it does frame quality in terms pertinent not just to industry, but also to collaborative work-flows in which volunteers and professionals both contribute. Indeed, in the project I am working on at the moment, the main working hypothesis is that maximum efficiency comes from arrangements that combine the work of trained transla-tors with the work of people trained in fields other than translation. I suspect that the subject-matter expertise of the latter group will complement the specific expertise of the trained translators.

Alongside the experiments, we are also in need of detailed observational studies. This is because, first, there are so many variables involved at different levels that only attentive qualitative research can really test their relations. Second, case studies are the only way to answer the more difficult questions about who (mis)understands whom, and who is ultimately profiting from new technologies. Whenever official discourse assures us that all is progress and all are benefitting, researchers should be able to base their critiques on more complex realities: we must seek and attempt to measure the communication gaps, before attempting to bridge them.

Why carry out this kind of research? For academics, one of the main reasons is training, as we have said: we have to design and teach courses where students become proficient in the skill sets required for work with these technologies. A related interest is then purely theoretical: thanks to these technologies, it is possible that the basic concept of "transla-tion" is changing, in ways that profoundly challenge several centuries of print culture. And a third interest should be critical: we at least have the liberty to ask a few hard questions that the technicians do not want to know about.

Of course, product designers will want us to tell them about the best kind of user interface and most efficient transla-tion practices, and companies will ask us how little they should pay translators, in the interests of ever greater productivity. Those sorts of questions can indeed be addressed in our research, and what we find should of course be applied in industry, as far as possible. Our own aims, though, need not be restricted to finding ways for other people to make more money.

V. POSTSCRIPT: IN SEARCH OF DEMOCRATIZATION

Herbert Marcuse [1] criticized technological rational-ity for reducing the individual to a crowd and imposing a fetish of efficiency. When submitting to ideologies of technol-ogy, individuals lose track of their own interests and accept the logic of the anonymous and uniform crowd. Marcuse's extreme pessimism was no doubt of its age; it belonged to an era when technologies were owned by the socially dominant classes, where they operated in the interests of those classes, and where the fetish of massification produced fascism. Should that critique of the crowd be applied to today's processes of "crowdsourcing"? Has efficiency really become the sole reigning rationality?

I would hope that interactive web-based technology allows for more collective partnerships than Marcuse envis-aged in 1941. I am even naïve enough to consider as valid the logics by which translators contribute to a privately owned database in order to reap the collective benefits of that data-base (such would still be the deal offered by Google Translator Toolkit, for example). Far more things are happen-ing than Marcuse had in mind. On the other hand, how successfully they are happening, and how far they can help the flourishing of the individual, remains to be seen. There are at least signs that the technologies may be both democratized and democratizing. We are moving toward an age where everyone will be able to translate, just as everyone can sing (some better than others, some for fun, some for money, but all with the basic capacity), and some are learning to sing to-gether, for wide audiences. The notes are not just in the right order – there are actual performances.

My argument here is clearly not that the academy should be training researchers for industry, although I have nothing against that happening. I am far more interested in breaking the barriers between research and practice in rather more creative ways. When my classes on translation technolo-

gies become sessions where trainees actually do research on their own performances [38], it is not hard to see how that sense of research can also enter what translators do when they test new tools, or when developers try out new ideas. In all these cases, on all those levels, individuals assume the freedom to find out what works best for them. If that can become the basic sense of research, then it might itself be the best way to make our technologies democratic.

REFERENCES

[1]   H. Marcuse, "Some Social Implications of Modern Technology" (1941) in A. Arato and E. Gebhardt, eds *The Essential Frankfurt School Reader*, New York: Continuum, 1982. 138-162, p. 153.

[2]   A. Way, "A critique of Statistical Machine Translation". *Linguistica Antwerpiensia* 8/2009, Special issue on Evaluation of Translation Technology. 17-41, p. 17.

[3]   L. E. Webb, *Advantages and Disadvantages of Translation Memory*, MA thesis. Monterey: Monterey Institute of International Studies, 1998-99.

[4]   K.-H Freigang, "Machine-aided translation". In M. Baker, ed. *Routledge Encyclopedia of Translation Studies*. London and New York: Routledge, 1998. 134-136.

[5]   S. O'Brien, "Practical experience of computer-aided translation tools in the localization industry". In L. Bowker, M. Cronin, D. Kenny and J. Pearson, eds *Unity in Diversity?: Current Trends in Translation Studies*. Manchester: St. Jerome, 1998. 115-22.

[6]   H. Somers, "Translation memory systems". In H. Somers, ed. *Computers and Translation: A Translator's Guide*. Amsterdam and Philadelphia: Benjamins, 2003. 31-47.

[7]   B. Dragsted, *Segmentation in Translation and Translation Memory Systems: An empirical investigation of cognitive segmentation and effects of integrating a TM system into the translation process.* Doctoral dissertation, Copenhagen Business School: Samfundslitteratur, 2004.

[8]   M. Yamada, "The effect of translation memory databases on productivity", in A. Pym, ed. *Translation research projects 3*. Tarragona: Intercultural Studies Group, 2011. 63-73.

[9]   C. Bédard, "Translation memory seeks sentence translator…". *Traduire*. 2000, p.186.

[10]   L. Bowker, "Productivity vs. Quality: A pilot study on the impact of translation memory systems", *Localisation Focus* 4(1), 2005. 13-20.

[11]   I. García, "Is Machine Translation Ready Yet?" *Target* 22 (1), 2010: 7-21.

[12]   F. Austermühl, "Training Translators to Localize". A. Pym et al., eds *Translation Technology and its Teaching (with much mention of localization)*, Tarragona: Intercultural Studies Group, 2006. 97-105.

[13]   C. Ribas, *Translation Memories as vehicles for error propagation. A pilot study*. Minor Dissertation. Tarragona: Intercultural Studies Group, Universitat Rovira i Virgili, 2007.

[14]   A. Pym, "Localizing Localization in Translator-Training Curricula", *Linguistica Antverpiensia* 33 (1999). 127-137.

[15]   A. Pym, "Localization and the Humanization of Technical Discourse", *Across Languages and Cultures* 4/2 (2003). 205-216.

[16]   A. Pym, "Localization and the Training of Linguistic Mediators for the Third Millennium", *Forum* 2(1) (2003). 125-135

[17]   A. Pym, *The Moving Text: Localization, Translation and Distribution*, Amsterdam & Philadelphia: John Benjamins, 2004.

[18]   A. Pym, "Translation as a Transaction Cost", *Meta* 40/4 (1995). 594-605.

[19]   L. Guerra Martínez, *Human Translation versus Machine Translation and Full Post-Editing of Raw Machine Translation Output*. Masters Dissertation. Dublin: Dublin City University, 2003.

[20]   A. Guerberof, *Productivity and quality in the post-editing of outputs from translation memories and machine translation: A pilot study*. Minor Dissertation. Tarragona. Universitat Rovira i Virgili, 2008.

[21]   A. Guerberof, "Productivity and quality in the post-editing of outputs from translation memories and machine translation". *Localisation Focus* 7(1) (2009). 11-21.

[22]   L. Thicke, "Improving MT results. A study". *Multilingual* Jan-Feb. 2011. 37-41, p. 40.

[23]   U. Reinke, "Evaluierung der linguistischen Leistungsfähigkeit von Translation Memory Systemen. Ein Erfahrungsbericht". *LDV-Forum* 16 (1999). 100-117.

[24]   C. Rico, "Evaluation Metrics for Translation Memories", *Language International* 12/6 (2000). 36-37.

[25]   F. Gow, *Metrics for Evaluating Translation Memory Software*. MA thesis. University of Ottawa, 2003.

[26]   R. Fiederer, and S. O'Brien, "Quality and machine translation: a realistic objective?" *Journal of Specialised Translation* 11 (2009). 52-74.

[27]   A. Birch et al., "Metrics for MT evaluation: evaluation reordering". *Machine Translation* 24 (1) (2010). 15-26.

[28]   K. Papineni, S. Roukos, T. Ward, and W. J. Zhu, W. J. 2002. "BLEU: a method for automatic evaluation of machine translation", *ACL-2002: 40th Annual meeting of the Association for Computational Linguistics*, 2002. 311–318.

[29]   C. Callison-Burch, et al. 2006. "Re-evaluating the role of BLEU in Machine Translation Research". In *Proceedings of EACL 2006* (11th Conference of the European Chapter of the Association of Computational Linguistics). Trento, Italy, 2006. 249-246.

[30]   F. Austermühl, *Electronic Tools for Translators*. Manchester: St. Jerome, 2001.

[31]   F. Austermühl, "Training Translators to Localize". A. Pym et al., eds *Translation Technology and its Teaching (with much mention of localization)*, Tarragona: Intercultural Studies Group, 2006. 97-105.

[32]   B. Esselink, *A Practical Guide to Localization*. Amsterdam and Philadelphia: John Benjamins, 2000.

[33]   G. Corpas Pastor, and M. J. Varela Salinas, eds *Entornos informáticos de la traducción profesional. Las memorias de traducción*. Granada: Atrio, 2003.

[34]   J. Torres del Rey, *La interfaz de la traducción. Formación de traductores y nuevas tecnologías*. Granada: Comares, 2005.

[35]   S. O'Brien, "Teaching Post-Editing: A Proposal for Course Content". *6th EAMT Workshop Teaching Machine Translation*- Manchester, 2002. 99–106.

[36]   B. Mossop, *Revising and Editing for Translators*. Manchester: St Jerome, 2001.

[37]   TAUS. *Post-editing in Practice*. Netherlands: TAUS BV, De Rijp, 2010.

[38]   A. Pym, "Using process studies in translator training. Self-discovery through lousy experiments", S. Göpferich, F. Alves & I. M. Mees, eds *Methodology, Technology and Innovation in Translation Process Research*. Copenhagen: Samfundslitteratur, 2009. 135-156.

[39]   A. Pym, "Translation research terms. A tentative glossary for moments of perplexity and dispute", Anthony Pym, ed. *Translation Research Projects* 3. Tarragona: Intercultural Studies Group, 2011. 75-110.

[40]   J. Holz-Mänttäri, *Translatorisches Handeln. Theorie und Methode*. Helsinki: Academiae Scientarum Fennicae, 1984.

[41]   C. Way, "Making theory reality. An example of interdisciplinary cooperation". In Georges Androulakis, ed. *Translating in the 21st Century. Trends and Prospects. Proceedings*. Thessaloniki: Aristotle University, 2003. 584-592.

[42]   T. Abekawa et al., "Community-based construction of draft and final translation corpus through a translation hosting site Minna no Hon'yaku (MNH)", in N. Calzolari et al. eds *Proceedings of the Seventh international conference on Language Resources and Evaluation* (LREC), Valletta: European Language Resources Association, 2010.

[43]   D. Lin, et al. "Composing Human and Machine Translation Services: Language Grid for Improving Localization Processes". In Nicoletta Calzolari et al., eds *Proceedings of the Seventh international conference on Language Resources and Evaluation* (LREC). CD. Valletta: European Language Resources Association, 2010.

[44]   J. Carson-Berndsen, H. Somers, C. Vogel, A. Way, "Integrated language technology as a part of next generation localisation," *Localisation Focus* 8(1) (2010). 53-66.

[45]   M. O'Hagan, "Fan translation Networks An Accidental Training Environment?", in John Kearns, ed. *Translator And Interpreter Training: Issues, Methods, Debates*. London: Continuum, 2008. 158-181.

[46] I. García, Ignacio, "Beyond Translation Memory: Computers and the professional translator". *The Journal of Specialised Translation* 12 (2009). 199-214.

[47] I. García, "The proper place of professionals (and non-professionals and machines) in web translation". *Tradumàtica* 8 (2010). unpaginated.

[48] A. Pym, "What technology does to translating", *Translation & Interpreting* 3(1) (2011). 1-9.

[49] P. Sánchez-Gijón, et al. "Traducción asistida, calidad y evaluación". N. Edo Marzá, P. Ordóñez López (eds) *El lenguaje de la ciencia y la tecnología.* Castelló de la Plana: Publicacions de la Universitat Jaume I, 2010.

[50] O. Torres-Hostench, et al.. "TRACE: measuring the impact of CAT tools on translated texts". Ll. Gea et al. eds *Linguistic and Translation Studies in Scientific Communication*, Peter Lang, 2010. 255-276.

[51] A. Chesterman. *Functional quality*, 2004. Video lecture at: http://www.facebook.com/video/video.php?v=1082863408296

[52] J. Katz, "Effability and Translation". In F. Guenthner and M. Guenthner-Reutter, eds *Meaning and Translation*. *Philosophical and Linguistic Approaches*. London: Duckworth, 1978. 191-234, p. 205.