Peter Sandrini, Marta García González (eds.)

Translation and Openness

EDITED VOLUME SERIES

Peter Sandrini, Marta García González (eds.)

Translation and Openness

Peter Sandrini
Institut für Translationswissenschaft, Universität Innsbruck

Marta García González
Facultade de Filoloxía e Tradución, Universidade de Vigo

Contents

Translation and Openness: an Introduction

Marta García González, Peter Sandrini
University of Vigo, Spain, University of Innsbruck, Austria

Openness includes removing barriers, taking away limits in order to allow access to and use of knowledge, content, data and software, as well as permitting sharing and collaboration. Openness implies transparency, something open is transparent for users, something that can be reproduced or verified, and something that doesn't conceal anything. When commercial interests are involved openness also means that these interests must be disclosed, they should be clear to users.

A trend towards a more collaborative society can generally be observed. Kennedy (2011), for example, describes three stages of social development, "corresponding very roughly to the first half of the 20th century (A), the latter half of the 20th century (B) and the beginning of the 21st century (C)" (Kennedy 2011: 6):

(A) Traditional	(B) Contemporary	(C) Emergent
rationalist economics	behavioural economics	knowledge society
rational	romantic	criticality
highly structured	neo-liberalism	distributed knowledge
top down	soft power	collaboration
centralisation	decentralisation	micro-agency
nationism/nationalism	globalisation	diversity
state power	localisation	public/private partnership
predictability	uncertainty	fuzziness/complexity
massproduction 'Fordism'	choice/market driven	mobility/flexibility
stratified society	less stratified society	multiple identities
collectivist cultures	individualism	participation

We cannot go into detail here, but the overall development tendency is "one from simplicity to complexity; from mono- to multi-dimensions; from structure to fluidity; from macro to micro" (Kennedy 2011: 7). With all these evolving trends, openness plays a key role, as a catalyst or facilitator. A knowledge society building upon distributed knowledge needs collaboration between the single actors, as well as access to knowledge for all people involved. Social roles shaped by diversity, flexibility and fuzziness are by definition open, and multiple identities, mobility and diversity inevitably presuppose an unprejudiced and open mindset.

The general notion of a free and open society gained a foothold in many branches of society: from ICT and technology with the concept of Free Software and the Digital Commons, law with open licenses such as the Creative Commons and the Copyleft licenses, pedagogy with the concept of Open Education and the sharing of educational resources (OER, MOOC), to public administration and the idea of Freedom of Information for public documents and processes put into practice by Open Government and Open Data, as well as research with the idea of Open Knowledge and Open Access. At the center of this trend stands the sharing of ideas and the vision of an open and free society and culture (e.g. Free Culture, Open Society Foundation).

Translation as social activity and Translation Studies (TS) as an academic discipline cannot elude those general tendencies. In fact, when we apply the characteristics of the emergent society (C) to translation we will see that many of these features are at the center of modern developments: participation and collaboration refer to participatory forms of translation (Cronin 2013; O'Hagan 2011) such as fansubbing, crowd translation, and all other types of voluntary translation listed in Desilets/van der Meer (2011: 29); multiple identities, flexibility, micro-agency lead us to the consolidation of the exciting branch of researching the sociological foundations of translation (Diaz-Fouces and Monzó 2010; Wolf and Fukari 2007); while the importance of knowledge, the role of the translator within a knowledge society, and distributed knowledge have been recognized widely in LSP translation (Budin and Lušicky 2014; Dam 2005) on the one hand, and in translation technology with the impact of the Internet on knowledge resources and translation data (Chan 2015), on the other hand.

Trying to define openness is not a trivial task: we may refer to the open definition website (opendefinition.org) where openness is defined in the context of open data, open content and open knowledge: "Knowledge is open if anyone is free to access, use, modify, and share it – subject, at most, to measures that preserve provenance and openness" (open definition, version 2.0); or refer to the concept of openness as used by the Free Software Foundation in describing free software and its use where they speak of four essential freedoms granted to users of free software:

- The freedom to run the program as you wish, for any purpose (freedom 0).
- The freedom to study how the program works, and change it so it does your computing as you wish (freedom 1). Access to the source code is a precondition for this.
- The freedom to redistribute copies so you can help your neighbor (freedom 2).

- The freedom to distribute copies of your modified versions to others (freedom 3). By doing this you can give the whole community a chance to benefit from your changes. Access to the source code is a pre-condition for this (gnu.org).

Free and open may not be used as synonyms, however. There was a long controversy going on between the Free Software Foundation and the Open Source Initiative about the very meaning of free and the ideology associated with it (Raymond 1999); eventually, it appeared that free means much more than open in the context of software, with the free software advocates insisting on freedom as the overall leitmotif and the more pragmatic Open Source followers emphasizing collaboration. Leaving aside ideological debates, we concentrate on using open and openness for the purpose of describing collaborative and free-availability behavior within translation.

Still, the concept of openness is a complex and multifaceted phenomenon touching many aspects of an activity or subject field. In particular, openness encompasses a range of topics (Educause 2009):

- Open standards and interoperability
- Open and community source software development
- Open access to research data
- Open scholarly communications
- Open access to, and open derivative use of, content.

For all these aspects, some initiatives or activities in translation can be found. According to a 2010 study (Gough 2011), 26% of translators explicitly endorse the "latest trends of sharing, openness and collaboration" (Gough 2011: 211) with more than 50% expressing a future commitment to these trends. While this study refers to practicing translators we may observe similar trends also in the academic world of translation studies.

Although in the field of translation and translation studies openness can be addressed from different perspectives, two lines of research have attracted particular attention in recent years, namely the study of open standards and formats in translation (Reineke 2005; Mata 2008) and the increasing move-ment towards open and collaborative forms of translation (O'Hagan 2011).

The use of open standards and formats in translation is relevant not only when connected to the actual behavior of professional translators (García González 2008), but also as a key element in translator training. As claimed by Mata (2008: 75-76), being familiar with the most common open standards and formats contributes to understand the importance and benefits of compatibility and interoperability of CAT tools and helps future translators to

informedly choose among the available tools based on their need and not only on the requirements of their customers.

Translation technology and the development of CAT tools is not any longer restricted to commercial providers as collaboratively organized open source projects are beginning to enter the desktop of professional translators and translator trainers. Translation memory systems, machine translation applications, text alignment tools, software localization programs, subtitling tools, text alignment and terminology tools, as well as translation management applications already exist as open source programs or free software. In many cases, users may even choose between two or more alternative packages. Openness in this respect not only facilitates access to such software applications or switching between different programs without any costs involved, it also enables users to contribute to these projects and to become part of a community.

Communities of users have evolved who regularly translate texts, documentation, film dialogues on a voluntary basis (O'Brien and Schäler 2010). These may be fan groups of television series or movies translating subtitles into many languages and sharing the translations on-line (fansubbing, fandubbing), fans of video games or users of free software who contribute to the projects by translating user interfaces or documentation material. Even companies with a large user base have begun to outsource the translation of their websites or on-line forums to their users (crowd-sourcing, user-generated translation) to economize on costs and time. These kind of translation done by lay people without any kind of specific training has become an object of study by the academic world with researchers investigating the efficiency and quality of their work, but also their impact on the professional world of translation (Olohan 2014; McDonough Dolmaya 2011 and 2012).

On the other hand, professional translators have begun to rediscover their ethical side and participate in voluntary translation work for NGOs. Some even have formed translation networks to deal with the large demand for translations by charitable bodies (e.g. Translators without Borders, The Rosetta Foundation, Mondo Lingua Initiative, Translators and Interpreters for Solidarity ECOS, Babels). On-line volunteer translators can be classified by their formal qualification, but also by their motivation and approach to translation, as done, for example, in Bey et al (2008: 136):

1. Mission-oriented translator communities: strongly-coordinated groups of volunteers involved in translating clearly defined sets of documents, mostly technical documentation.

2. Subject-oriented translator network communities: individual translators who translate on-line documents such as news, analyses, and reports and make translations available on personal or group web pages.

In many cases of volunteer translation we may observe a trend to "demonetization and deprofessionalization of translation" (Olohan 2014: 18) which is why openness is strongly opposed by many professional translators who strive to earn their living from translation. In view of these persisting and increasing trends, however, a lock-down or defensive attitude should give way to a more viable diversification and differentiation of translation as an activity.

The advantages of openness have been recognized also in the world of academia where the growing costs for journal subscriptions and publishers have begun to raise barriers for research. It is clear that research can thrive only when based upon other research, and thus, unrestricted on-line access to scholarly research is a necessary requirement. In March 2015, UNESCO launched its Open Access Curriculum, a set of manuals to facilitate capacity building of library and information professionals and researchers, as part of its Strategy on open access to scientific information and research. And we may observe a growing trend in academic translation journals to publish in an open access format as described in two contributions in this volume, so that open access to scholarly literature is beginning to gain a foothold also in translation studies.

Openness includes open access to, and open derivative use of content, in our case of translations. Translation technology and translation data allow the re-use of previously done translations on a broad scale, as implemented by statistical machine translation and translation memory systems. In the professional world of translation this has raised a number of questions, such as, for example, who owns a translation memory, how much price reduction can be applied in cases of a translation match of whatever percentage from a client-supplied translation memory, or what compensation should be paid when the translator is providing her translation memory to the client. It seems that in this case we are witnessing a conflict about who will be the ultimate beneficiary of economies of scale in translation. There is no doubt, however, that open content and open access to translation resources is important, especially in the context of official translations. Translations done by official institutions entirely financed from public funds should be made publicly available, not just as translated texts but also in the form of translation memories wherever available. Open access to translation data, thus, can be a part of an Open Government and Open Data strategy.

Contributions to this volume review some of the above referred topics, such as FOSS for translators and the training of translators with FOSS applications, or the open access to scholarly literature but also cover some other topics connected to the study of openness as it is quality, both quality of FOSS for translators and quality of volunteer and collaborative translations. Full coverage of all topics regarding openness in translation is beyond an anthology like this, the whole concept of openness is simply too varied and challenging.

Nevertheless, the volume falls into three thematic sections: the first and most substantial part deals with the concept of openness in ICT (open data, open tools, open computer systems, and quality evaluation of open software), the middle part is concerned with translators training and the use of open software, and the last part discusses openness in academia on the basis of the concepts of Digital Scholarship and the 'Scientist 2.0'.

The volume opens with a critical discussion of the concepts of openness and closedness/proprietariness as they relate to the assemblages of data, knowledge and information that result from the practice of professional translation. **Philipp Neubauer** underlines the fact that neither concept can be considered as existing in a vacuum, and that both need to be seen to play out against the background of social and technological change in society in general and a notable power differential between the suppliers and providers of translation services in particular. Special attention is to be drawn to the emergence of unintended consequences which may accompany processes of both "open sourcing" and appropriation of said resources.

Cristian Lakó then describes a methodology which takes freely available open tools on the web to set up a list of most used keywords relevant for the target audience. Thus, the profiling of the reader is no longer constructed on rather random data but on hard statistical evidence, and the target text, especially websites and other marketing oriented texts, is more likely to be found by the web-users of the target market, thus facilitating organic B2C communication.

In the third contribution, **Peter Sandrini** investigates why and how the free operating system GNU/Linux is suitable as a platform for multilingual text production and translation by outlining the rationale behind their development and their historical evolution. He presents several specific initiatives and examples of GNU/Linux based open desktop systems for translators and discusses potential reasons why a wider adoption in the translation community has not yet taken place.

Potential users of open-source translation technologies face the daunting task of considering the available options and selecting the one that better

satisfies their needs. **Silvia Flórez** and **Amparo Alcina** propose a quality model for the evaluation of open-source translation technologies going beyond software product evaluation and including aspects of the communities and processes that sustain development projects. Evaluation instruments and results are publicly available on-line.

Evaluation is also at the center of the following contribution: after a short over-view of the phases and results of the research project *Creación dunha plataforma docente GNU/LINUX para a formación de tradutores – localizadores de software – subtituladores*, funded by Xunta de Galiza, within the framework of programme Incite, **Maite Veiga Díaz** and **Marta García González** describe a particular research effort devoted to the testing of the usability of free and open-source translation memory managers and text aligners with different types of texts, and their applicability to translator training. This represents a smooth transition to the next topic of the volume, namely openness in a didactic context and specifically, translators training.

Approaches to process-oriented translator training can be optimized using freeware and FOSS screen recording technology. Screen recording technol-ogy captures all activity that transpires on-screen over the course of task completion in the form of a video that can be analyzed in a retrospective fashion for purposes of enhancing problem and problem-solving awareness, among other things. In addition to describing how to best utilize various fea-tures inherent to freeware and FOSS screen recording applications, **Eric Angelone** also presents a series of concrete learning activities as a ground-work guide for process-oriented training.

Adrià Martín-Mor, Ramon Piqué Huerta and **Pilar Sánchez-Gijón** from the Tradumàtica group show how openness is becoming a key concept in translation through a case in point: the collaboration between the Tradumàtica Masters (Translation Technologies) and the Public Knowledge Project (PKP) to localise their academic software (Open Journal Systems and Open Monograph Press) into Spanish and Catalan. This intersection between openness, trans-lators training and open access publication options brings us to the last the-matic division of the book which is openness in research and the academia.

The most important research tools, archives, libraries, research centers and universities make use of the central features of the web represented by the opportunity to save time and costs with connecting a wide variety of con-tent through linking. These emerge also as advantages in scientific publishing where such trends seem to be able to revolutionize research and scientific publishing activity. While open publishing and transparency seem to find more followers in the natural sciences, they are still far from being broadly accepted in the humanities, especially within the philologies. In his contribution, **Marco**

Agnetta describes the concept of a "Scientist 2.0" and investigates current opinions about open access that can be relevant for the self-conception of a future translatology by identifying strengths and weaknesses in positive and negative attitudes towards open access.

In the last contribution to the volume, **Peter Sandrini** gives an overview over digital scholarship in translation studies by examining publication methods and academic evaluation approaches where open initiatives and commercial activities confront each other. The author makes a plea for openness since more openness could very well foster the discipline of translation studies as a whole and move it towards a more unified and collaborative field of study.

Authors and editors have teamed up to put together a list of bibliographical references that aims at covering the different topics of openness and translation, a rather difficult task since such a compilation can never be exhaustive nor complete. The resulting list under the heading "Further Literature and Useful Readings" includes 179 references which may be subdivided into four sections:

- open tools (in translation) (82)
- open access (in translation studies) (7)
- open standards and formats (in translation) (9)
- open and collaborative translation (83)

Each reference is tagged with one or multiple keywords from this classification so that readers may identify which topic is covered. The digital version of the list of references (see web page at http://www.petersandrini.net/transopen.html) in BibTeX format allows for an automatic extraction of references according to a specific subfield; for this volume, however, an alphabetical arrangement was chosen because multiple categorizations would not be possible in the printed medium.

While openness regarding translation technology, or the development and adoption of open standards and formats may represent a rather clear-cut subject, for different reasons this is not the case with open and collaborative translation and open access in translation studies. Open and collaborative translation represents a very heterogeneous subject field including such diverse topics as community translation, user-generated translation, volunteer translation, crowd-sourcing of translation, and fan translation, fansubbing, fandubs, scanlators, etc. (for a detailed discussion of these concepts, their definitions and overlapping areas see O'Hagan 2011: 13-16). Moreover, this field of study has generated great interest among researchers and a lot of relevant

publications exist. Since this does not constitute the main topic of this volume, nor is it the goal of this compilation of references to cover all aspects of collaborative translation, we concentrated on the aspect of openness within this broad range of topics.

For a different reason, open access in translation studies represents another problematic classification. Much has been published about open access in general, but, unfortunately, very little related specifically to openness and open access in translation studies. Compiling a list of references, thus, represents a tedious task.

A chapter with short biographical notes on authors and a keyword index close the book.

We hope that readers will find this volume informative and that they will make use of the references given in order to further develop ideas and thoughts expressed in the contributions. As editors of this volume we are convinced that thinking about openness and implementing openness in our attitudes and actions have considerable bearing on our conception of ourselves as translators or researchers. Openness indeed questions the very role of translated texts, multilingual translation resources, the ethics of translators, their professional behavior, the self-conception of academics and researchers, as well as the role and availability of research results in society. Furthermore, openness challenges traditional commercial models both for professional translation and for academic publishing. It therefore constitutes one of the most stimulating challenges that the world of professional translation and translation studies have yet faced.

Acknowledgements

References

Bey, Y., Boitet, C. and Kageura, K. (2008) The TRANSBey Prototype: An Online Collaborative Wiki-based CAT Environment for Volunteer Translators. In Yuste Rodrigo, E. (ed.)

Topics in Language Resources for Translation and Localisation. Amsterdam: John Benjamins, 135-150.

Budin, G. and Lušicky, V. (2014) Languages for Special Purposes in a Multilingual, Transcultural World. Proceedings of the 19th European Symposium on Languages for Special Purposes, 8-10 July 2013, Vienna, Austria. Available at: http://lsp2013.univie.ac.at/ proceedings [Accessed 10 August 2015].

Chan, Sin-wai (ed.) (2015): The Routledge encyclopedia of translation technology. *London: Routledge.*

Cronin, M. (2013) *Translation in the digital age*. London: Routledge.

Currie, C. (2009) What Is Openness, Anyway? EDUCAUSE Quarterly. Available at: http://www.educause.edu/ero/article/what-openness-anyway [Accessed 10 August 2015].

Dam, H. V. (ed.) (2005) Knowledge systems and translation. Berlin: Mouton de Gruyter.

Désilets, A. and van der Meer, J. (2011) Co-creating a repository of best-practices for collaborative translation. In Linguistica Antverpiensia 10, 27-46. Available at: https://lans-tts.ua.ac.be/index.php/LANS-TTS/article/view/276 [Accessed 10 August 2015].

Diaz-Fouces, O. and Monzó, E. (2010) What would a sociology applied to translation be like? In MonTI 2, 9-18. Available at: http://rua.ua.es/dspace/bitstream/10045/16432/1/MonTI_2_01.pdf [Accessed 18 September 2015].

EDUCAUSE Review (2009) EDUCAUSE Values: Openness, January/February 2009. Available at: http://www.educause.edu/ero/article/educause-values-openness [Accessed 10 August 2015].

Garcia, M. (2008) Free software for translators. Is the market ready for a change, In Diaz Fouces, O. and García, M. (eds.) Traducir (con) software libre. Granada: Comares, 9-31.

Gough, J. (2011) An empirical study of professional translator's attitudes, use and aware-ness of Web 2.0 technologies, and implications for the adoption of emerging techno-logies and trends, In O'Hagan, M. (ed.) Translation as a Social Activity – Community Translation 2.0, Linguistica Antverpiensia 10/2011, 195-217.

Kennedy, C. (2011) Challenges for language policy, language and development. In Coleman, H. (ed.) Dreams and realities: developing countries and the English Language. London: British Council. Available at: https://www.teachingenglish.org.uk/sites/teacheng/files/Z413%20EDB%20Section02_0.pdf [Accessed 10 August 2015].

Mata Pastor, M. (2008) Formatos libres en traducción y localización. In Diaz Fouces. O. and García, M. (eds.) Traducir (con) software libre. Granada: Comares, 75-122.

McDonough Dolmaya, J. (2011) Wikipedia survey IV (motivations). Some thoughts on translation research and teaching. Available at: http://mcdonough-dolmaya.ca/2011/08/24/wikipedia-survey-iv-motivations/ [Accessed 10 August 2015].

McDonough Dolmaya, Julie (2012): Analyzing the crowdsourcing model and its impact on public perceptions of translation. The Translator 18, no. 2, 167-191.

O'Hagan, M. (2011) Community Translation: Translation as a social activity and its possible consequences in the advent of Web 2.0 and beyond. In Linguistica Antverpiensia. 10, 1-10.

O'Brien, S.; Schäler, R. (2010) Next generation translation and localization: Users are taking charge. Paper presented at Translating and the computer 32, London, November, 18-19. Available at http://doras.dcu.ie/16695/1/Paper_6.pdf [Accessed 10 August 2015].

Olohan, M. (2014) Why do you translate? Motivation to volunteer and TED translation, Translation Studies, 7:1, 17-33.

Raymond, E. (1999) The Cathedral and the Bazaar: Musings on Linux and Open Source by an Accidental Revolutionary. Boston: O'Reilly Media.

Reineke, D. (2005) XMK en la traducción, en Reineke, D. (ed.) Traducción y localización: mercado, gestión y tecnologías. Las Palmas: Anoart Ediciones, 285-315.

Wolf, M. and Fukari, A. (eds.) (2007) Constructing a Sociology of Translation. Amsterdam: John Benjamins Publishing Company.

Unforeseen Consequences: Big Data and the Language Industry

Philipp B. Neubauer
Independent Researcher

1 Introduction

There are some long-term consequences of technological change that affect specific areas of social experience in ways that cannot in a direct or straight-forward way be deducted from the intentions of the actors who are involved in bringing them about. For this reason, they are of considerable importance to social scientists and there is a long tradition of studying these so-called un-foreseen or unintended consequences. Merton (1936) is considered to be the first to have set down systematic observations on the topic (Dietz 2004). Two key points of his observations are that unforeseen consequences need not be identified with axiologically negative effects (Merton 1936: 895) and that it need "not [be] assumed that in fact social action always involves clear-cut, ex-plicit purpose" (ibid: 896/897). It is however safe to assume that the construc-tion of a scenario that plausibly charts the context in which the unforeseen consequences are situated would be beneficial to their study and evaluation. This is the stated purpose of the present article. It is intended to provide some impulses for the study of unforeseen consequences of technological change – of course, our speculative/heuristic method can only produce hypotheses whose evaluation would then fall into the purview of empirical sociology and/or translation studies research, the disciplines which need to come up with designs for representative surveys – both to sociologically oriented re-searchers in translation studies (and particularly to those pursuing approaches based on the sociology of professions (Stichweh 2005), e. g Diaz-Fouces and Monzó 2010; Sela-Sheffy 2011: 11) as well as to anyone in-terested in the broader field of technology assessment (Kalverkämper 1998: 12). This is to be achieved by charting some correlations between tendencies of the language services market and the context of industrial processes in-volving statistical machine translation (SMT) and post-editing (PE) within the bigger picture of the big data paradigm as it takes shape in the language in-dustry on the one hand and the conceivable consequences this may have for the perception and economic position of translation professionals on the other hand.

Given that many of the emergent effects can be seen as "foreseen"/ intended – or at least as assented to and accepted – on the part of large supply-side language industry players, there are already impressionistic

studies or personal commentaries on their impact on the translating profession (Rudavin 2009; Katan 2011) or critiques that focus on the influence of technology use on conceptions of translation equivalence and vice versa (Nogueira de Andrade Stupiello 2008). If one aims to bring the unforeseen and unintended into focus, one might look at them from the perspective of the advocates of free, libre and open source software and open access content, as this draws attention to the seeming paradox that e. g. de-professionalization might occur as a side effect of justified demands for accountability (Sandrini 2013; Mayer-Schönberger and Cukier 2013: 116), the democratic strife for access to education and freedom of information (Heylighen 2007) or simply as epiphenomena contingent on technological development. The epistemic opportunity in this regard lies in contrasting and synthesizing the perspectives of translators/post-editors and open source advocates precisely because there seems to be so little overlap between these subcultures, if one extrapolates from the current prevalence and uptake of FLOSS translation tools (García Gonzalez 2008).

Part of this synthesis will consist in arriving at a "sociological glimpse" (Diaz-Fouces/Monzó 2010: 10) which accounts for the sentiments and impressions of individual actors in the translation market. Then we will briefly expound on the ethos of open source and open access for the purpose of distinguishing, from this point of view, intended consequences from unintended/unforeseen ones. Following this, we shall introduce some more detailed observations on the technological developments driving structural change on the part of language industry suppliers:

1. Big data as a general technological trend towards the aggregation and algorithmic parsing of ever larger amounts of data; this general trend can serve as a template for interpreting developments in the translation services market by analogy.

   a) Statistical Machine Translation (SMT), which represents the application of statistical algorithms to large repositories of translation data, e. g. such composed of translation memories (TM), on-line bitexts and parallel texts and especially the so-called open data, which public institutions disclose or release to the general public (Sandrini 2013). Another factor driving the growth of accessible translation data can be seen in the traction gained by open formats for data interchange (ibid.) which (at least in theory) facilitate the aggregation of data by ensuring its uniform structural presentation.

   b) Post-editing (PE), by which we primarily refer to the rewriting of machine translation output in order to achieve results that are comparable to human translation, this is the subclass of "full post-

editing" (Allen 2003: 306). Within the scope of this article, this is the only relevant type as our argument depends on the commensurability with fully human (intellectual) translation. The output of PE activity can subsequently be added to the machine translation corpora used as its starting point. PE itself can be organized in the form of crowdsourcing (compare Fédération Internationale des Traducteurs (FIT) 2015) or it can be cast as a new way of professional translating, albeit one fraught with new challenges. This is reflected in the emergence of formal training courses in post-editing for which certification is available, for instance at the language service provider SDL plc. (2015a) or the industry association TAUS (2015).

Concluding the article, we will co-ordinate the insights into the technical workings of SMT/PE with the sociological glimpse obtained in the first section, which shall lead to an evaluation of the present trend in conjunction with a forecast of what there might be to come.

2 A Sociological Glimpse of the Language Industry

Here, the situation regarding the progressive automation of the workplace in general may serve as a starting point; it is noteworthy that in recent years this seems to have begun to penetrate to professions that would previously have been considered impervious to automation. According to an article published in Wired Magazine (Dormehl 2015) which quotes research by the University of Oxford conducted in 2013, approximately 47% of all jobs are predicted to be cut due to automation over the course of the next 20 years – the exact scope of the study in terms of industry and geographic scope was not amplified on; while this trend has been around since the dawn of the industrial revolution in the 19th century, its new quality seems to be that now, "white-collar professions involving a high level of training are just as likely to be displaced by software [...] because once-untouchable fields such as law and medicine include specialisms that are vulnerable to automation: medical diagnosis, the drafting of contracts and comparison of trademarks can be better carried out by a computer than by human beings" (ibid.). The researchers who published the study saw the reason for this in the fact that the subdivision of larger work processes into ever smaller series of actions, which has greatly facilitated the automation of "cognitive work".

Although this prognosis with its more general scope does not make any specific mention of the language industry or the market for translation services, the scenario seems to resonate with some observer's laments about the degradation in pay, prestige and working conditions that seem to prevail in

this area. Often, their blame is laid on technical innovation and/or economic developments.

Where technical innovation is concerned, the reason for the downward spiral is attributed to changes in perception regarding the translator and his or her task brought about by machine translation and translation memory technologies. One example for this is the critique articulated by Nogueira de Andrade Stupiello (2008), whose views shall be briefly summarized here. Contrary to the creed of functionalism, translators in highly automated environments are no longer seen as responsible for the semantic rendering of the target text, but are seen to be merely tasked with cosmetic changes to the semi-automatically generated output, which – as folk wisdom would have it – is already semantically complete and fully equivalent of the source. Hence, the focus is on minor flaws, details that the machine could not successfully "recover". According to the critic, this perspective itself is not new, but follows from the tradition of translation technology and is already manifest in the conventions of translation memory use. Here, leverage is paramount even if the pre-translated segments do not fit their new context and thus any retranslation of existing matches due to textual concerns is neither desired nor remunerated. Nogueira de Andrade Stupiello (2008) thinks that the reasons for the prevalence of these attitudes can be found in the ever-shorter production cycles for translations, the need to cut cost and the "urgency of communication" under the pressures of globalization and the information age, which must eventually lead to lowered expectations regarding linguistic quality. At the end of the day, all that seems to matter is to somehow grasp the gist of a foreign language text.

Rudavin's (2009) observations, by contrast, are formulated from a personal and practice-oriented perspective. He is concerned especially with the market situation of freelance translators, whereby the focus is less on technology assessment or the profession's image in a stricter sense and more on the underlying structure of the language industry and its tendencies as a business sector. He observes that as such, the language industry cannot be viewed in isolation from its larger economic context and its actor's financial incentives. In this regard, he also names "globalization" as the key driver, besides "market consolidation" and technical progress. The interrelation of the latter two is of special interest here: as global ITC networks facilitate the coordination of international multilingual projects, there emerges a market for projects which, due to time constraints, scale and the number of languages required are only manageable by the largest language service providers, actors whom Rudavin calls "translation corporations". In some cases, these happen to be the very same corporations who also act as vendors of proprietary CAT tools that provide the workflow/process infrastructure by

which translation tasks devolve to smaller subcontracting agencies and ultimately the freelance translators. According to Rudavin, the "translation corporations" (which remain unnamed) already have a strong foothold in the market; the 30 largest vendors together are said hold a market share of 20% at an annual growth rate of 20-50%. If this tendency were to continue, a likely consequence would be the formation of an oligopoly.

3 Big Data, Open Source and Open Data

This is the initial scenario that we shall assume for the critique of the unforeseen/unintended consequences of the use of open and public data and open source technology in a for-profit translation context, since a starting hypothesis about the priorities and interests of industry actors is necessary for deducting intentions and contrasting them with the unintended/unforeseen consequences of their social actions. Before this can be attempted, there remain the enabling technological conditions to be explored.

3.1   Big Data

As it shall be seen, the big data paradigm is central to the success of the method of statistical machine translation while certain forms of openness can be seen to constitute necessary preconditions for the application of the big data paradigm to the language industry. There is hitherto no complete intensional definition of big data, however, two essential properties indicative of this state of social and technical development can be identified: on the one hand, there is a steady increase in the quantity of digital data as the digitization of ever more areas of human experience progresses; on the other, there is an emergent qualitative change of the area itself which follows the utilization of the data in its respective context. This latter is what Mayer-Schönberger and Cukier (2013: 6) assert to be the defining attribute of big data:

> [D]ata has begun to accumulate to the point where something new and special is taking place. [...] The quantitative change has led to a qualitative one. The sciences like astronomy and genomics, which first experienced the explosion in the 2000s, coined the term "big data". […] There is no rigorous definition of big data. [...] One way to think about the issue today [...] is this: big data refers to things one can do at a large scale that cannot be done at a smaller one, to extract new insights or create new forms of value, in ways that change markets, organizations, the relationship between citizens and governments, and more.

If it is assumed that SMT (with or without downstream PE) constitutes a new mode of value creation for the language industry which has the potential to

disrupt markets and production processes, the question remains where exactly the mass ("big") data fueling the SMT engines are sourced from and how they are exploited or ultimately monetized.

3.2   Open Source

One possibility for obtaining the mass data is to rely on open sources, whereby this statement can be confusing as the data in question need not be licensed as "open source" as in "free, libre and open source", but need only be publicly and unrestrictedly accessible, as in "open-source intelligence" (Wikipedia contributors 2015c, Open-source intelligence) – The Open Source model (Heylighen 2007a) itself follows a principle similar to that of "communalism", which is at work in the organization of science (Merton 1988: 680); thus, the Mertonian concepts used to describe scientific organization should be reasonably continuous with this new context. Nevertheless, such data can and does include "free and open" licensed sources in a stricter sense. According to FOLDOC (2012: Open Source), this is the intention behind Open Source as a model of software licensing and distribution:

> A method and philosophy for software licensing and distribution designed to encourage use and improvement of software written by volunteers by ensuring that anyone can copy the source code and modify it freely.

This concept, which reflects a denotation of unlimited redistribution and modification, is not limited to software products, but applies to other immaterial goods as well. Insofar as a strict separation of formal language texts and digital natural language data and audiovisual material is tenable (compare Touretzky 2001), it has been designated either Open Access (Heylighen 2007) or Open Content (Gunn 2008) where it relates to the latter. Analogous to the family of open source software licenses, a few licensing models for Open Content can be distinguished from the published content itself. According to Gunn (2008), the "Creative Commons" (CC) and "Free Document" (FDL) licensing models can be cited as examples of explicitly free and open licenses for publishing. The intentions motivating Open Data initiatives which also include open translation data (compare Sandrini 2013: 33) can be seen to vary somewhat from this theme. Here one might distinguish explicitly open from public data, with the latter satisfying the criterion of de-facto open access without necessarily being meant for free redistribution and modification.

3.3   Open Data

True open data originate with the public sector and government institutions (Sandrini 2013: 33); they are often released to the general public because

public institutions can rarely do more than merely administer the data on behalf of their constituencies for want of resources and expertise (Mayer-Schönberger and Cukier 2013: 116).

An example for open translation data can be found in those published by the European Union (ibid.) who also hope to advance their own SMT program in this fashion. Translation data of the UN have been published in the context of the "Corpora Commons" initiative, also with the explicit aim of furthering SMT research (Gunn 2008). These two examples concern open data in the stricter sense (compare Mayer-Schönberger and Cukier 2013: 38); patents and trademarks which must by decree be published in several languages (Pariser 2011) might serve as yet another example.

The development of Google Translate, currently perhaps the most prototypical phrase-based statistical machine translation system, exemplifies the conflation of open and public data in the training of SMT engines; besides the actual open data aggregates described above, public data comprising practically all translation data of the world wide web have been leveraged for its training. Among this, there has been some with contentious legal status, as the utilization of translations from the Google Books project shows – see "Authors Guild, Inc. v. Google, Inc." (Wikipedia contributors 2015) (Pariser 2011).

While the for-profit use of true open data is (at least in general under-standing) in line with the intentions of their providers, the same treatment of merely public data constitutes a gray area at very minimum. This might also be applicable to some extent to proprietary translation data held by language service providers, provided that they meet two conditions: firstly, they need to be fungible, i. e. come in a structurally open (interchange) format (Sandrini 2013: 33) and secondly, they need to be scrambled by technical means in order to circumvent some intellectual property laws that would otherwise apply to the data in aggregate (Zetzsche 2005); this at least holds inside the German jurisdiction (Cruse 2014) and shows that determining the status of such data is difficult to begin with. Once the conditions are met, these data might also be treated as public.

3.4   Distinguishing Public Data and Open Source Software

While these considerations reference the relationship of SMT and data, open source software is also directly and indirectly relevant to developments in SMT. For one, free and open source SMT software and components immediately lower the barrier for SMT research (Lopez 2008: 3), while a more indirect consequence can be discerned in the diversity of ideas, actors and projects and the flat hierarchies of open source development (Heylighen

2007) which favor rapid evolution. Even though our focus lies on data as the main driver of SMT uptake, these factors might be of interest in the assessment of any unforeseen consequences stemming from the FLOSS paradigm itself – a conceivable case in point is the use of free and open SMT systems, e.g. MOSES (2015) on the part of language service providers. Though this appears a plausible scenario, there now seems to be (to the best of my knowledge) no economically significant use of this or similar systems – however, if any such use were modeled on the patterns described here, they would qualify as cases for the study of the unforeseen effects of FLOSS products.

Considering that data is the key component, it is for now safe to neglect the impact of the actual licensing model of SMT software on the scenario to be devised. Its basis lies in the construction of the relationship between the availability of data to fuel data-driven semi-automatic production processes on the one and the structure of these processes, i. e. how language workers interface with machine output, on the other hand.

4 Machine Translation and Post-editing

Research into machine translation has been around since the advent of electronic computers in the 1940s (Ping 1998: 162). Historically, the area has seen its ups and downs, the former marked by irrational exuberance triggered by an overestimation of the impact of advances in memory capacity and computing power on machine translation capabilities, the latter by the subsequent disenchantment caused by the evaluation of the actual results delivered by predominantly rule-based historical machine translation systems (Weizenbaum 1976: 186). Such tendencies are still extant, however, the premise seems to have changed with the shift towards big data/statistical processing; here, it is plausible to assert that increasing "processor speed, random access memory size, secondary storage, and grid computing" will indeed contribute to the improved performance of machine translation systems (Lopez 2008: 3) because such performance would be based on a larger throughput of data (i. e. larger amounts parsed) to begin with.

However, this article is not intended be an in-depth review of the history, functional principles and limitations of the machine translation systems themselves; we merely draw on these to elucidate on its argument. The focus is more on current tendencies in the actual deployment of SMT systems that can be linked to both big data and open data than on their history or technical details. The following figure shows a breakdown of MT systems by the fundamental strategy used to create the semblance of a "translation"

performance on chunks of natural language input and thus a "pseudo-translation" (Torrens, cited in Wilss 1996a: 212).
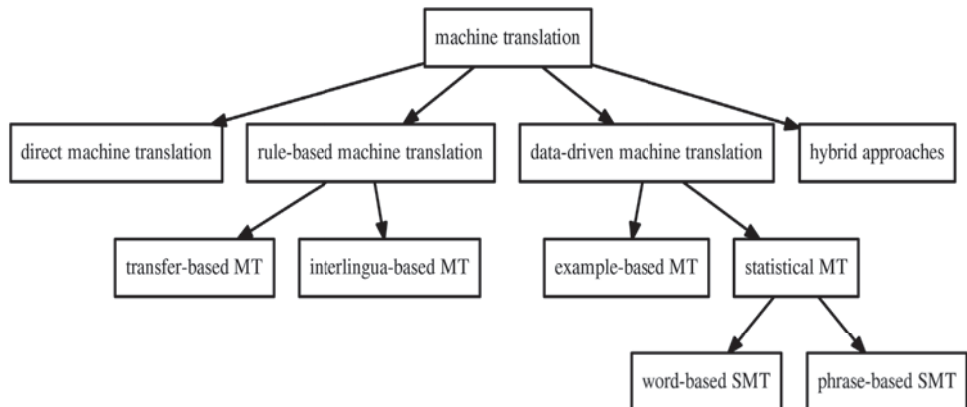

Figure 1: Taxonomy of machine translation architectures. Based on: Labaka et al. 2007; Lopez 2008; Eberle 2008; Gupta 2012; Okpor 2014.

If one completely disregards both the historical strategy of direct machine translation and any hybrid approaches there remain two fundamentally different strategies of MT, the rule-based and the data-driven. The rule-based model aims at generating a pseudo-translation by means of a pre-encoded linguistic and grammatical rule set for a generative transfer of L1 to L2. The statistical model relies on parsing large quantities of data for the probability of translation equivalence and thus constitutes the kind of technology that might benefit significantly from a quantitative hike in the available data. Here, we can discern the potential for the conversion of quantity to quality that Mayer-Schönberger and Cukier have envisioned.

4.1   Statistical Machine Translation

This potential lies in the reliance on statistical correlations between L1 and L2 renderings of chunks or phrases (in the case of the currently prevalent phrase-based SMTS, Lopez 2008: 9) rather than on explicit grammatical rules for the generation of a pseudo-translation. The linguistic material for analysis resides in parallel corpora (i. e. aligned translation data) parsed by the SMT algorithm. Unlike the rule-based model, the machine makes no attempt at emulating human interpretation or reconstructing the semantics of the source text (Ping 1998: 163-164). It does however appear to demonstrate "machine learning" (Lopez 2008: 1) in the sense described here:

A control system acts when there is a discrepancy between what it senses (sensory signal) and what it is supposed to sense or would like to sense (reference). The connections that matter are those of certain activities in the system's repertoire with the changes they provoke in certain sensory perturbations. A mechanical feedback device that replaces us in a given task is a crystallized piece of experiential learning. It is the materialization of an if-then rule that has been inductively derived from experience by the designer (Glasersfeld 1981).

What the machine "likes to sense" in this case is the larger probability of a given L1 phraseme having been translated by L2 phraseme X, as opposed to phrasemes Y, Z and so on. This figure shows what remains at the end of the mapping process:
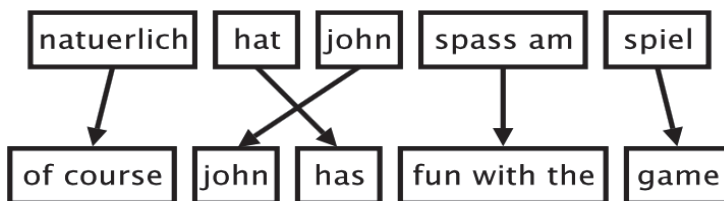

Figure 2: A phrase-based SMT model; Koehn (2010).

This however also serves to illustrate that the machine will only be capable of providing a "plausible" pseudo-translation if the search space for such probabilities is large enough, both in terms of finding positive correlations for the largest possible amount of L1 phrasemes and in terms of eliminating relatively unlikely candidate phrases; as the search space thus equals the corpus of phrase pairs "known" to the algorithm, it becomes clear why SMT performance is linked closely to corpus size and (alignment) quality (Arnold 2003: 139; Lopez 2008: 1; Labaka et al. 2007).

It also shows that the approach of so "guessing" the probability of a phrase to appear in a certain slot regardless of its semantic function is a far cry from the (always contested) idea of artificial intelligence as aiming "to simulate human intelligence as it manifests itself in the understanding of all reality, concrete or abstract, with which human beings are confronted [... b]y means of entirely automatic processes" (Wijnands 1993: 166). If one tries, for the sake of the argument, to imagine the pseudo-translation process as performed by a human, one might think of someone who is neither a speaker of L1 nor L2 in the process of assembling fragments of "fuzzy matches" from a translation memory system, guided only by their optical resemblance to

character strings which appear in L2 texts. Insofar as reading the pseudo-translation can be said to have caused someone to "understand" its intended message, this would have been a function of the database/corpus having contained very similar phrasal material, which in turn would only have been likely (read: probable) if the search space was very large indeed; this is why SMT is considered a big data application (de Palma 2013).

That this event is even possible constitutes the previously mentioned "new quality from quantity"; as recently as 12 years ago, the scarcity of data had been seen as a severe limitation of the statistical approach to machine translation (Arnold 2003: 139). Now, the increasing availability of open and public translation data have made this a non-issue, at least for some language combinations. Predictably, this increase in the volume of data has translated into better quality pseudo-translations (Scholtes 2010), to the extent that the technology has now attracted the interest of language service providers (Rex 2013) and the largest technology players (Herranz 2014) alike. Even if the quality of the output of the free (of charge) web translation offers (e. g. Google Translate) is scarcely good enough for integration into professional translation workflows, this need not be the case for proprietary engines offered by language service providers like SDL ("BeGlobal", SDL plc. 2015b) which have been trained on well-aligned and often industry-specific input data.

4.2   Post-editing

However, to reiterate our argument, neither a large statistical search space nor a cleanly aligned MT corpus can in and of themselves grant the SMT engine the capability to translate in the sense of producing something that actually equals a human translation in form and function. It lacks the crucial element of "intelligence", however one likes to define it (Wilss 1996b; Weizenbaum 1976: 186-187). Whether or not one believes that the original meaning of the source text can somehow be "recovered" from the phrase salad resulting from SMT or whether one asserts that it takes an act of interpretation of the pseudo-translation relative to the source in order to arrive at a semantically viable reading of any pseudo-translation that does not by chance resemble a natural language utterance (which need not bear any semantic relationship to the source language's) is moot with regard to this statement.

To my mind, this is about the pinnacle of the "translation performance" that current systems are capable of. That the public and scientific interest in machine translation research has never completely waned despite this might be explained by venturing that linguistic utterances do not "contain" any intrinsic meaning, but that any meaning is synthesized by the recipients' fitting

them into their experiential world. It is this act which provides considerable leeway for the benevolent interpretation of pseudo-translation as well as that of any other speech act (especially those in written language) (Berman 2013: 2-4; von Glasersfeld 1999).

If SMT technology is to be employed for the creation of value on the basis of big data, the missing ingredient needs to be added downstream, at a later stage of the production process. This stage is called post-editing (PE); it involves the use of human labor to impose potential meaning by rewriting/ reordering the SMT pseudo-translation. In principle, this understanding does not significantly deviate from the definition of post-editing as the "the correc-tion of machine translation output by human linguists/editors" (Veale and Way, cited in Allen 2003: 297). It seems likely that the literature contains many more variations on this theme.

Any of these might however be open to criticism, both from the vantage point of translation theory and from that of statistical machine translation tech-nology itself. On the one hand, the notion of "correction" reflects the some-what naive view of natural language criticized by Nogueira de Andrade Stupiello (2008), namely that which maintains that essential meaning (to the extent that this is believed to inhere in the source) has already been recov-ered by the SMTS and that the segment would only need to be polished by removing minor errors (e. g. non-agreement of suffices, superfluous or missing words and other artifacts of alignment). However, it should now have become clear that this essentially contradicts the premise of an a-semantic and non-interpretative mode of pseudo-translation generation. Insofar as a meaning is read into the signage of the segment by the post-editor or subse-quent interpreter, its emergence is owed to the intervention of the person's consciousness and their ability to interpret language within considerable tolerances – it has clearly not been actively recovered by the machine. As the term "segment" in this context suggests, the primary locus of "meaning recov-ery" is – in line with the prevalent design logic of current translation editor soft-ware – the micro-linguistic level of the sentence or below, where accidental matches are far more probable than on the macro-linguistic level of the com-plete text. Here, the chances for these to occur should be astronomically small, which is probably why the impact of SMT on texts hardly seems to fea-ture in considerations of SMT capabilities. Granting the possibility of "lucky" selections on the segment level and minimal human intervention with the out-put of well-trained engines, the translation performance proper as it is per-ceived by the final recipient needs ultimately be enacted by the human post-editor, not the engine, which can't (and isn't designed to) provide it.

Having stated this, there is also the aspect of SMT economy to consider. While it is always possible to replace an inviable pseudo-translations with a completely new translation, this is certainly not the best solution in terms of leverage, considering that the post-edited output is not only there to serve the immediate need of the translation customer, but that it should ultimately return to the SMT corpus in order to enlarge its search space (i. e. the range of phrase variety covered) and so to guarantee future leverage for more plausible pseudo-translations.

"Leverage" in this sense can be understood as analogous to the use of this term in the context of translation memories, i. e. better leverage is achieved by (re-)using as many of the original SMT suggestions as possible in order to closely match similar input in the future; depending of the quality of the SMT corpus used, it is easy to see how this goal competes with that of efficiently imposing potential meaning. Incorporating both these competing goals into the PE strategy can be seen as a challenge notably absent from conventional human translation.

Hence, the capability for reconciling and balancing the human and machine demands of the task – i. e. the demand for communicative meaning and readability on the one, the demand for uniformity and future leverage on the other hand – is the distinguishing quality of post-editing when compared to translating. However, with regard to the more standard qualities demanded in commercial translation (correctness, speed, and cost), there is no question of "either ... or"; the additional challenges of post-editing simply add to the overall requirements. This translates into cumulative difficulty, as post-editing has the goal of translating more text faster. The PE additionally faces the challenge of submitting more text to QA procedures, etc. in even less time. Post-editing, which in this way differs from purely human translation both in terms of quality and of quantity, can thus appear a task that "anyone can do" (Pym 2013: 489) only at the most superficial of enquiries.

5 A Tentative Scenario for the Translation Market

To conclude this line of enquiry, it now remains to relate the aspects of underlying technology to the impressions of our "sociological glimpse". The connecting elements are both the status of the translating profession as an income-generating factor (or, on the reverse, the decreasing rates which are a hallmark of de-professionalization) and the competition between translation workers with differing qualification profiles (compare Monzó 2011). The heart of the matter is that post-editing as an occupational activity does not seem to belong to any recognized profession which in turn would lend it the pedigree

correlated with higher remuneration (Fuchs-Heinritz et al. 1995: 521). The following statements are indicative of this observation:

- Pym (2013: 491) understands post-editing as an area associated with "technical communication" but notes that efforts at professionalizing this discipline tend to lag far behind those already undertaken for translation and interpreting;
- Allen (2003: 298-299) observed that, at least at the time of his writing, hard-and-fast criteria to certify the qualification of post-editors were lacking; recent efforts to formalize this qualification, like those already mentioned, might remedy this in the short term but will never convey the professional pedigree of a full university degree program.

Given that the self-reported status of translators in a recent study (Katan 2011: 77-78) was relatively low – respondents stated that is was largely comparable to that of a "secretary" – and that tendencies of de-professionalization are already under investigation (ibid 66) in this field, the key danger is to my mind that due to the nature of the process, crucial human capabilities are either accidentally misattributed to the SMT engines or deliberately misrepresented. If so, the likely consequence is a further erosion of the professional recognition of translators/PEs, aggravated further by clients being isolated from the translation/localization process by multiple layers of large language service provider's corporate bureaucracies, two factors which are very likely to coincide, especially when these middle-men are vendors of language services and translation technology/SMT products at the same time.

The peril for the translation/PE practitioner lies less in falling victim to an actual deskilling, insofar as this is defined as a "reduced utilization of [... and] partial or complete devaluation of existing scholastic/academic, professional or vocational qualifications" (Fuchs-Heinritz et al. 1995: 135, my translation), as should have emerged from the present discussion. It lies in the loss of (or rather the failure to attain) the professional standing which secures expert status and monetary perks for the members of the more prototypical professions (Katan 2011: 70).

From this apparent de-professionalization results a change in the structure of competition in the market; when linguistic competence is devalued or no longer counts as a distinguishing professional qualification (Pym 2013: 489), a situation may emerge in which translation/PE professionals will have to compete against those whose qualifications are either completely different or those whose (source-)language competence might be significantly worse than is acceptable for professional translators (Katan 2011: 71). This larger competitive field may ultimately lead to further downward pressure on prices and/or

the exclusion from business opportunities of those who can't (or won't) compete under these circumstances.

This is likely to affect projects which are very demanding in terms of subject competence, e. g. specialized translations relating to law or medicine (where there perhaps might already be a possibility for semi-automatically generating the source text) as well as those where the expectations in terms of visibility and linguistic quality are very modest, e. g. "F.A.Q" sections for consumer products and the like.

This conclusion readily agrees with Rudavin's (2009) observation that subject specialists with a second language have recently been preferred over those who are (only) professional translators for complex assignments in the above fields. Add to this the observation that "[...] you often have no constant need to look at the foreign language [...] for some low-quality purposes, you may have no need to know any foreign language at all, if and when you know the subject matter very well" (Pym 2013: 489) and it should be easy to see how a combination of SMT/PE-capabilities and extant labor market tenden-cies might generate a synergy to that effect. This means that the growth of translation data (e. g. when already-dominant LSPs manage to appropriate large high-quality corpora for specific domains) which contributes to the recognizability/interpretability of pseudo-translations coincides with the automation of certain professions that may lead to the simultaneous "release" of a significant numbers of workers. The displacement of specialized translators by SMT-augmented multilingual specialists for the field in question would at least be a conceivable outcome. This scenario is not without a parallel in already existing situations where markets/fields of competence overlap (Katan 2011: 73); yet, the aspect of combined technological and social change holds the potential for bringing about a new, unforeseen quality in this phenomenon.

It seems even more likely when we approach the market for low-end translation services. As specialist knowledge does not matter here, there might even be a market for anonymous crowdsourcing workflows. Since the professional association Fédération Internationale des Traducteurs (FIT) (2015) has already published a position paper outlining the method of crowdsourcing, we will not amplify on this matter here; our assertion is the emergence of a scenario akin to that outlined for high-complexity projects, only with an aggravated tendency towards "lowest-bid market economics" (Muzii, cited in Katan 2011: 66). Translation workers will thus compete via pricing rather than competence/qualification. Between the high and the low end of the market, a visual breakdown of the projected scenario in relation to current practices might look like this:
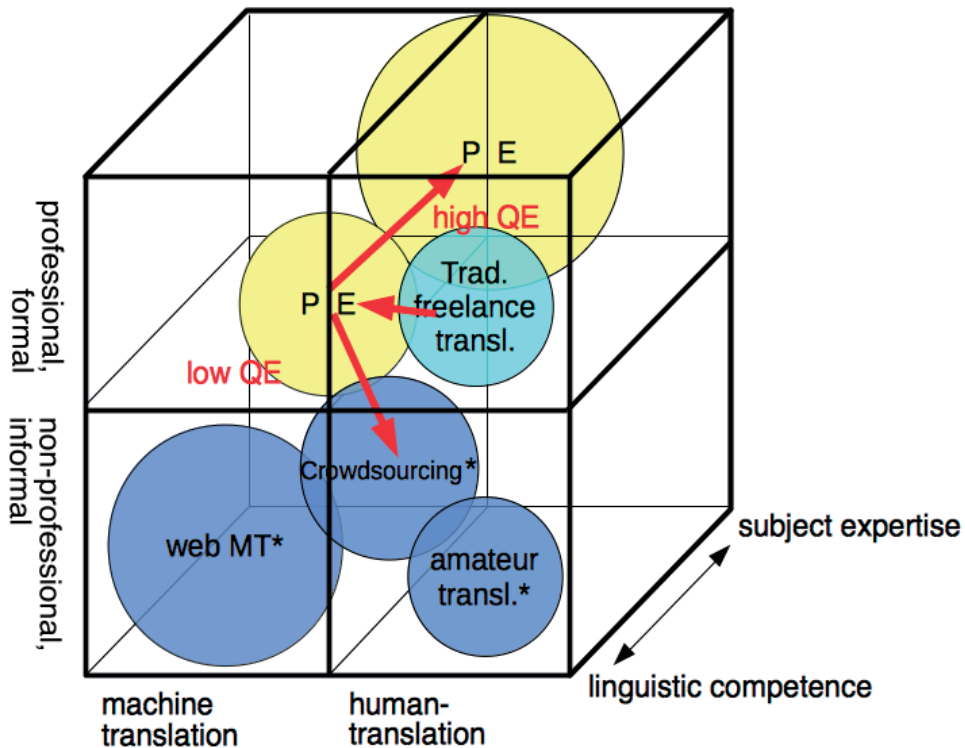
Figure 3: Intellectual translation vs. post-editing; the depth of specialized knowledge cannot be determined for activities marked with an asterisk.

For this we use a modified priority matrix with an added dimension of depth (linguistic competence vs. subject expertise). The matrix is inscribed with Venn diagrams showing any overlap between types of activities. Traditional (freelance) translating entails working a diverse portfolio of both classical translation and PE, highly specialized and general jobs, etc. It thus occupies a median position. In contrast to this, there is the noted drift towards the "back" of diagram in PE with high expectations in terms of quality (QE). Low-QE post-editing overlaps crowdsourcing in the lower right quadrant, which – due its black-box nature – may overlap with and introduce both raw machine translation from web engines and unrevised amateur human translation.

6 Outlook and Concluding Remarks

While it is conceivable that the scenario we have envisioned is likely both foreseen and intended on the part of language service providers, it is a cogent question to ask whether these consequences have been foreseen – or could have been foreseen – by any of those who have contributed to creating the basis of this economy of human/machine translation: institutional decision makers releasing open data to the public, developers of algorithms and (open source) software, academics concerned with basic research in fields like linguistics, mathematics, computer science and many more. From their vantage point, the unforeseen consequences of the growth of both open and public translation data can best be attributed to Merton's category of "chance consequences", "occasioned by the interplay of forces and circumstances which are so complex and numerous that prediction of them is quite beyond our reach" (Merton 1936: 899-900), owing to the fact that either of these endeavors seem remote from the translation services market and that there does seem to be an element of the co-incidence of a number of disparate developments involved. Nevertheless, we have managed to construct a scenario "on the ground" by identifying and connecting some of these forces and circumstances for the purpose of discussing their interplay; they are:

- the increasing automation of cognitive work,
- the role attempts at value creation through the combined use of big data resources and statistical machine learning algorithms play in this,
- the shifting expectations of translation consumers and language service providers brought about by market consolidation, globalization and the progress of certain technologies,
- the accelerated technical change through community-driven and open scientific research and software development modeled on analogous patterns,
- the economic rationalization of workflows through the combined use of human and machine resources, which gives rise to the practice of post-editing.

The most noteworthy paradox that rears its head here is that the unforeseen consequences of de-professionalization and falling proceeds from translating – even if they appear to be results of a very indirect causality – glaringly contradict the stated intention of the push to open translation data, namely to "enhance the perceived value of translation and to elevate the status of translators as a professional group" (Sandrini 2013: 33, my translation). This leaves the question of the final lesson learned from tackling the phenomenon. What the present author is paid for post-edited words is exactly half of what

the same customer is willing to pay for "new words" of a conventional human translation. If this is in any way indicative of an emergent industry trend would again need to be established by means of a representative study.

If one belongs to a group that is put a disadvantage by current developments, it is certainly tempting to feel a nostalgic longing for the "old days" of closed-off, guild-like professions and to renounce the open and collaborative mode of work which threatens to dissolve inherited privilege, even if scholars in the sociology of professions point out that the traditional professions are losing their former social and economic traction anyway (Stichweh 2005) and if one takes into account that privilege and closure in this sense have been considered an unfair advantage over laymen since the days of Adam Smith. Keen (2008) can be named as an example for this reactionary outlook on contemporary technology and culture. It seems however rather doubtful that such musings can provide any positive impulses for engaging with the present professional practice or for shaping the future of translation as a business.

They also miss the essential point. As already suggested, the true peril seems to consist in too little openness and transparency rather than too much. It would be a function of cumulative advantages – this is a concept from the sociology of science (Sismondo 2010: 39-40) which generalizes Merton's "Matthew effect" (Merton 1968: 58; Merton 1988: 609); it might be understood as a form of positive feedback which leads to "inequalities [...that] appear to result from self-augmenting processes" (Merton 1988: 617). These effects, initially observed in scientific careers, also form a sub-category of unintended consequences (Merton 1988: 615). Apparently not limited to science, they can be observed in similar social fields, e. g. open source software development, where Heylighen (2007) observed a "'rich get richer' dynamics [negatively affecting] equally valuable, competing projects [which,] because of random fluctuations or sequence effects, may fail to get the critical mass necessary to 'take off'". Such cumulative advantages are garnered by the "translation corporations" as a consequence of their growth and economies of scale that coincide with an environment characterized by an accelerated de-professionalization of language services in combination with a distorted perception of human/computer PE/SMT processes. Either is a consequence contingent on the big data phenomenon and some mutual interdependence can be ascribed to them.

Providing that storing larger quantities of data opens new qualitative paths for its commercial exploitation, vendors of SMT systems might start off by training their engines on open translation data and expand their reach by re-training them with data for other languages and domains as they flow back from their normal translation/PE operations. As the recognizability/ interpreta-

bility of pseudo-translations improves with rising corpus size, it will become possible for them to shunt existing customers from human translation to SMT/PE-based processes, whereby the deal can be sweetened for the consumer by passing some of the cost reductions on to them. This might create a virtuous circle (from the vendor's vantage point) as more data is funneled back into the engine, more customers are attracted and the vendor's economic clout increases. Consequently, they will find themselves in a position where they are increasingly capable of dictating (lower) translation purchasing prices and of squeezing competitors out of the market.

Any such (hypothetical) companies are practically doomed to appear as "free riders" from the vantage point of the institutions and communities that contribute technology and data in accordance with the open source ethos (Heylighen 2007): industry preferences for proprietary licensing, vendor lock-in and draconian non-disclosure agreements all but preclude any data, knowledge or technological improvement from being given back to the communities and general public. Such would be the working of a "ratchet effect" that allows the free flow of open and public resources into proprietary systems, but not the other way around.

*Figure 4: The "Mechanical Turk", a 19th century make-believe chess automaton.
Source: Wikipedia contributors 2015d*

Translators/post-editors would likely be affected in a different way. Here, the gap for exploitation lies in the representation of machine capabilities and their actual inability to produce more than pseudo-translations. Even if it can be assumed that no reputable language service provider would ever try to conceal this fact from their customers, downplaying it for marketing purposes would not be considered unethical by many. The human PE, the real engine of the process who ultimately bears the responsibility for the usefulness of the product – its fitness for the purpose of human communication – is blotted out from the perception of the translation consumers and thus enacts a role that begins to resemble that of the operator working in the interior of the "Turk" (Wikipedia contributors 2015d, The Turk) who helps create and maintain the illusion of an autonomously playing chess automaton by lending his or her capability to the "machine".

Ironically, this will reinforce the impression of the "non-human, technical [...] habitus" ascribed to translating (Katan 2011: 78) and executives' imputed opinion of translators as "human-mechanical revenue generating machines" (Rudavin 2009) – with all the perfectly foreseeable socio-economic conse-quences this is likely to have for the practitioners themselves.

Due to the complexity of the interplay of macro-social and technological forces that bring about similar developments, a public debate of the desirable and undesirable consequences of data-driven technologies in general is likely to benefit not only translation businesses, professional associations and translation studies as an academic discipline, but also society at large. If we fail to practice technology assessment in time, we are at peril of being overwhelmed by unforeseen consequences in the long run.

References

Allen, J. (2003) Post-editing. In Somers, H. (ed.) Computers and translation: a translator's guide. Amsterdam: John Benjamins, 297-317. Available at: https://books.google.de/books?id=a4W7lWgCqYoC [Accessed 18 September 2015].

Arnold, D. (2003) Why translating is difficult for computers. In H. Somers (ed.) Computers and translation: a translator's guide. Amsterdam: John Benjamins, 119-142. Available at: https://books.google.de/books?id=a4W7lWgCqYoC [Accessed 18 September 2015].

Berman, J. J. (2013) Principles of big data: preparing, sharing, and analyzing complex information. Oxford: Newnes. Available at: https://books.google.de/books?id=gEho0DI8a2kC [Accessed 18 September 2015].

Cruse, A. (2014) Besitzansprüche – Urheberrecht und elektronische Datensammlungen. *MDÜ* 14.3, 10-15.

de Palma, D. A. (2013) Big Data Comes to the Translation Sector. Common Sense Advisory Blogs. Available at: http://www.commonsenseadvisory.com/default.aspx?Contenttype=ArticleDetAD&tabID=63&Aid=3025&moduleId=390 [Accessed 18 September 2015].

Diaz-Fouces, O. and Monzó, E. (2010) What would sociology applied to translation be like? In MonTI 2., 9-18. Available at: http://rua.ua.es/dspace/bitstream/10045/16432/1/MonTI_2_01.pdf [Accessed 18 September 2015].

Dietz, H. (2004) Unbeabsichtigte Folgen – Hauptbegriff der Soziologie oder verzichtbares Konzept? Zeitschrift für Soziologie 33.1, 48-61. Available at: http://www.zfs-online.org/index.php/zfs/article/viewFile/1154/691 [Accessed 18 September 2015].

Dormehl, L. (2015) Your job automated. *Wired Magazine* (UK Edition) 01.15, 126-133. Available at: http://www.wired.co.uk/magazine/archive/2015/01/features/your-job-automated [Accessed 18 September 2015].

Eberle, K. (2008) Integration von regel- und statistikbasierten Methoden in der Maschinellen Übersetzung. *Journal for Language Technology and Computational Linguistics* 23.2, 37-70. Available at: http://www.jlcl.org/2009_Heft3/kurt_eberle.pdf [Accessed 18 September 2015].

Fédération Internationale des Traducteurs (FIT) (2015) FIT Position Statement on Crowdsourcing of Translation, Interpreting and Terminology Services. Online position paper. Available at: http://www.fit-ift.org/wp-content/uploads/2015/04/Crowd-EN.pdf [Accessed 18 September 2015].

FOLDOC Contributors (2012) Free On-line Dictionary of Computing (FOLDOC). Online lexical database. Howe, D. (ed.) Available at: http://foldoc.org/ [Accessed 18 September 2015].

Fuchs-Heinritz, W. et al. (1995) Lexikon zur Soziologie, 3., völlig neu bearbeitete und erweiterte Auflage. 3rd edition. Opladen: Westdeutscher Verlag. Available at: https://books.google.de/books?id=wSieBgAAQBAJ [Accessed 18 September 2015].

García Gonzalez, M. (2008) Free Software for translators: is the market ready for a change? In Diaz-Fouces, O. and García Gonzaléz, M. (eds.) Traducir (con) software libre. Granada: Comares, 3–31.

Glasersfeld, E. von (1981) Feedback, Induction, and Epistemology. In Applied systems and cybernetics. Lasker, G.E. (ed.) Vol. 2. New York: Pergamon Press, 712-719. Available at: http://www.univie.ac.at/constructivism/EvG/papers/069.pdf [Accessed 18 September 2015].

Glasersfeld, E. von (1999) How Do We Mean? A Constructivist Sketch of Semantics. In Cybernetics and Human Knowing 6.1, 9-16. Available at: http://www.univie.ac.at/constructivism/EvG/ papers/221.pdf [Accessed 18 September 2015].

Gunn, Allen (2008) Open Translation Tools: Disruptive Potential to Broaden Access to Knowledge. Report prepared for the Open Society Institute. Open Society Institute. Available at: http://aspirationtech.org/files/AspirationOpenTranslationTools.pdf.

Gupta, S. (2012) A survey of data-driven machine translation. Available at: http://www.cfilt.iitb.ac.in/resources/surveys/MT-Literature%20Survey-2012-Somya.pdf [Accessed 18 September 2015].

Herranz, M. (2014) Twitter, eBay, Facebook. Big data companies want to own machine translation. Pangeanic blog post. Available at: http://blog.pangeanic.com/2014/08/10/twitter-ebay-facebook-big-data-companies-want-to-own-machine-translation/# [Accessed 18 September 2015].

Heylighen, F. (2007) Why is open access development so successful? Stigmergic organization and the economics of information. In Open Source Jahrbuch. Lutterbeck,

B., Bärwolf, M. and Gehring, R.A. (eds.) Lehmanns Media. Available at: http://pespmc1.vub.ac.be/Papers/OpenSourceStigmergy.pdf [Accessed 18 September 2015].

Kalverkämper, H. (1998) 1. Fach und Fachwissen [Subject and subject knowledge]. In Hoffmann, L. and Kalverkämper, H. (eds.) Fachsprachen – Ein internationales Handbuch zur Fachsprachenforschung [Languages for special purposes – An international handbook of special-language and terminology research] Vol. 1. Berlin: de Gruyer, 1-24.

Katan, D. (2011) Occupation or profession, A survey of the translators' world. In Sela-Sheffy, R. and Shlesinger, M. (eds.) Identity and Status in the Translational Professions. Amsterdam: John Benjamins, 67-87. Available at: https://books.google.de/books?id=KbZxAAAAQBAJ [Accessed 18 September 2015].

Keen, A. (2008) The Cult of the Amateur: How blogs, MySpace, YouTube, and the rest of today's user-generated media are destroying our economy, our culture, and our values. New York: Crown Business. Available at: https://books.google.de/books?id=Z59TDBx1U2UC [Accessed 18 September 2015].

Koehn, P. (2010) Chapter 5: Phrase-Based Models. Available at: http://www.statmt.org/book/slides/05-phrase-based-models.pdf [Accessed 18 September 2015].

Labaka, G. et al. (2007) Comparing rule-based and data-driven approaches to Spanish-to-Basque machine translation. In Proceedings of the MT Summit XI. European Association for Machine Translation. Available at: http://doras.dcu.ie/15228/1/LabakaEtAl_summit_07.pdf [Accessed 18 September 2015].

Lopez, A. (2008) Statistical machine translation. In ACM Computing Surveys (CSUR) 40.3, 8. Available at: https://alopez.github.io/papers/survey.pdf [Accessed 18 September 2015].

Mayer-Schönberger, V. and Cukier, K. (2013) Big data: A revolution that will transform how we live, work, and think. Boston: Houghton Mifflin Harcourt. Available at: https://books.google.de/books?id=uy4lh-WehhIC [Accessed 18 September 2015].

Merton, R. (1936) The Unanticipated Consequences of Purposive Social Action. In American Sociological Review 1.6. 894-904. Available at: http://users.ipfw.edu/dilts/E%20306%20Readings/The%20Unanticipated%20Consequences%20of%20Purposive%20Social%20Action.pdf [Accessed 18 September 2015].

Merton, R. (1968) The Matthew Effect in Science. In Science 159.3810. Key concept source, 56-63. Available at: http://www.garfield.library.upenn.edu/merton/matthew1.pdf.

Merton, R. (1988) The Matthew Effect in Science, II – Cumulative Advantage and the Symbolism of Intellectual Property. In ISIS 79. Key concept source, 606-623. Available at: http://garfield.library.upenn.edu/merton/matthewii.pdf [Accessed 18 September 2015].

Monzó, E. (2011) Legal and translational occupations in Spain, Regulations and specialization in jurisdictional struggles. In Sela-Sheffy, R. and Shlesinger, M. (eds.) Identity and Status in the Translational Professions. Amsterdam: John Benjamins, 11-30. Available at: https://books.google.de/books?id=KbZxAAAAQBAJ [Accessed 18 September 2015].

MOSES Project (2015) Welcome to Moses! (Statistical machine translation system). Community website. Available at: http://www.statmt.org/moses/ [Accessed 18 September 2015].

Nogueira de Andrade Stupiello, É. (2008) Ethical Implications of Translation Technologies. In Translation Journal 12.1. No longer available.

Okpor, M.D. (2014) Machine translation approaches: issues and challenges. In IJCSI International Journal of Computer Science Issues 11.5, 159–165. Available at: http://www.ijcsi.org/papers/IJCSI-11-5-2-159-165.pdf [Accessed 18 September 2015].

Pariser, E. (2011) The Filter Bubble, how the new personalized web is changing what we read and how we think. New York: Penguin. Available at: https://books.google.de/books?id=wcalrOI1YbQC [Accessed 18 September 2015].

Ping, K. (1998) Machine Translation. In Baker, M.; Saldanha, G. (eds.) Routledge Encyclopedia of Translation Studies. 2nd Edition. London: Routledge, 162-170.

Pym, A. (2013) Translation Skill-Sets in a Machine Translation Age. In Meta 58.3, 487-503.

Rex, M. jr. (2013) Exploring the Intersection of Big Data and Machine Translation. TAUS blog post. Available at: https://www.taus.net/think-tank/articles/translate-articles/exploring-the-intersection-of-big-data-and-machine-translation [Accessed 18 September 2015].

Rudavin, O. (2009) Current trends in the translation industry and what they mean to us all of us. In Baur, W. et. al (eds.) Übersetzen in die Zukunft, Herausforderungen der Globalisierung für Übersetzer und Dolmetscher, Tagungsband der Internationalen Fachkonferenz des Bundesverbandes der Übersetzer und Dolmetscher e.V. (BDÜ). Vol. 32. Schriften des BDÜ. Berlln BDÜ, 69–75.

Sandrini, P. (2013) Open Translation Data – Die gesellschaftliche Funktion der Übersetzungsdaten. In Mayer, F. and Nord, B. (eds.) Aus Tradition in die Zukunft: Festschrift für Christiane Nord. Berlln Frank & Timme, 27–37.

Scholtes, J.C. (2010) Machine Translation that Works, Finally! Here is why and how... eDiscovery and Information Risk Management, Blog. Available at: https://zylab.wordpress.com/2010/03/31/machine-translation-that-works-finally-here-is-why-and-how.../ [Accessed 18 September 2015].

SDL plc. (2015a) Post-Editing Machine Translation Certification. Corporate website. Available at: http://www.translationzone.com/learning/training/post-editing-machine-translation/ [Accessed 18 September 2015].

SDL plc. (2015b) SDL BeGlobal, Cloud-based machine translation for high-volume, fast communication. Corporate website. Available at: http://www.sdl.com/cxc/language/machine-translation/beglobal/ [Accessed 18 September 2015].

Sela-Sheffy, R. (2011) Introduction: Identity and Status in the Translational Professions. In Sela-Sheffy, R.; Shlesinger, M. (eds.) Identity and Status in the Translational Professions. Amsterdam: John Benjamins, 1–9. Available at: https://books.google.de/books?id=KbZxAAAAQBAJ [Accessed 18 September 2015].

Sismondo, S. (2010) An introduction to Science and Technology Studies. 2nd Edition. London: Wiley-Blackwell.

Stichweh, R. (2005) Die Soziologie der Professionen. Working paper, Universität Bonn, Abteilung Demokratieforschung. Available at: http://www.fiw.uni-bonn.de/

demokratieforschung/personen/stichweh/pdfs/38_die-soziologie-der-professionen-_2_.pdf [Accessed 18 September 2015].

TAUS (2015) Post-editing Course. Association website. Available at: https://postedit.taus.net/ post-edit/training-certification [Accessed 18 September 2015].

Touretzky, D. S. (2001) Viewpoint: Free speech rights for programmers. In Communications of the ACM 44.8. Extended online version, 23–25. doi: 10.1145/381641.381651. Available at: http://www.cs.cmu.edu/~dst/DeCSS/Gallery/ cacm-viewpoint.html [Accessed 18 September 2015].

Weizenbaum, J. (1976) Computer power and human reason – from judgment to calculation. New York: W.H. Freeman.

Wijnands, P. (1993) Terminology vs. Artificial Intelligence. In Sonneveld, H.; Loening, K. (eds.) Terminology – Applications in interdisciplinary communication. Amsterdam: John Benjamins, 165-180.

Wikipedia (2015b) Authors Guild, Inc. v. Google, Inc. In Wikipedia, The Free Encyclopedia. Wikipedia contributors (eds.) San Francisco, CA: Wikimedia Foundation Inc. Available at: https://en.wikipedia.org/wiki/Authors_Guild,_Inc._v._Google,_Inc [Accessed 18 September 2015].

Wikipedia (2015c) Open-source Intelligence. In Wikipedia, The Free Encyclopedia. Wikipedia contributors (eds.) San Francisco, CA: Wikimedia Foundation Inc. Available at: https://en.wikipedia.org/wiki/Open-source_intelligence [Accessed 18 September 2015].

Wikipedia (2015d) The Turk. In Wikipedia, The Free Encyclopedia. Wikipedia contributors (eds.) San Francisco: Wikimedia Foundation Inc. Available at: https://en.wikipedia.org/wiki/The_Turk [Accessed 18 September 2015].

Wilss, W. (1996a) Knowledge and Skills in Translator Behaviour. Amsterdam: John Benjamins.

Wilss, W. (1996b) Translation as intelligent behaviour. In Somers, H. (ed.) Terminology, LSP, and translation: Studies in language engineering in honour of Juan C. Sager. Amsterdam: John Benjamins, 161-168.

Zetzsche, J. (2005) TM Marketplace White Paper, Sharing Translation Memory Data Aligned from Third-Party Documents: Legal Considerations. Available at: http://www.tmmarketplace.com/whitepapers/align.pdf [Accessed 18 September 2015].

Search Engines and Related Open Tools for Establishing a Term Base

Cristian Lakó
Petru Maior University, Tg. Mureș, Romania

1 Introduction

In this paper we speak of openness in translation in the context of collecting and curating a terminology database for the purpose of translating on-line content in the case of multilingual websites. Whereas openness in translation is often considered from the perspective of the (on-line) tools employed (free vs. paid) or from the point of view of the translatum producers (community enthusiasts vs. professionals), we suggest using open and on-line tools for determining a term base, as a pre-editing translation process. A term base is required for consistency all over the translated content of a website and based on user input in search engines. Search engines such as Google, Bing, and Yahoo collect user input and make it available for on-line marketing purposes as keywords. Such keywords, in this case considered as central words in a text, can be regarded as translation suggestions to be used in a target text (TT). Translation based on this approach is often referred to as SEO (Search engine optimization) translation and SEO localization and make the process of opting for "the right translation" be grounded on statistical data; therefore translation is no longer a decision-making process. A similar concept to SEO translation is international SEO.

Also, as a pre-editing translation method, this approach corroborates Nord's instrumental translation (2005), and Eugene Nida's receptor-oriented theory (Dimitriu 2009: 26) by accurately establishing a common linguistic context between the text producer and the potential target readers. The usage of keywords determines the context of the TT, further emphasizing that translation can function as "an independent message transmitting instrument in a new communicative action in the target culture" (Nord 2005: 81). From a strictly linguistic point of view, Nord's definition of instrumental translation, can be also referred to as part of the localization process as we will see later on. From the perspective of localization, researched keywords can represent the local mix or locale (seen in this case as a group of users with similar interests) and they can also be used to profile the potential search engine users. By choosing the appropriate keywords (see long-tail keywords below) most search engine users can become receivers and not just addressees (see Nord's distinction – 1997: 22).

Using keywords as the starting point in the translation process is justified when considering that the most efficient way of on-line marketing is through web pages (see Figure 1). The main component of web pages is content, especially searchable textual content indexed by search engines. This is a solid argument to build a term base founded on keyword research.

Sphere size indicates level of usage

Level of Effectiveness

Customer Reviews
Webpages
Webinars/ Webcasts
Whitepapers
Microsites
E-newsletters
Articles
Case Studies
E-books
Blogs
Press Releases
Online Video
Images
Social Media (other than blogs)
Mobile Content
Digital Magazines
Podcasts

Degree of Difficulty
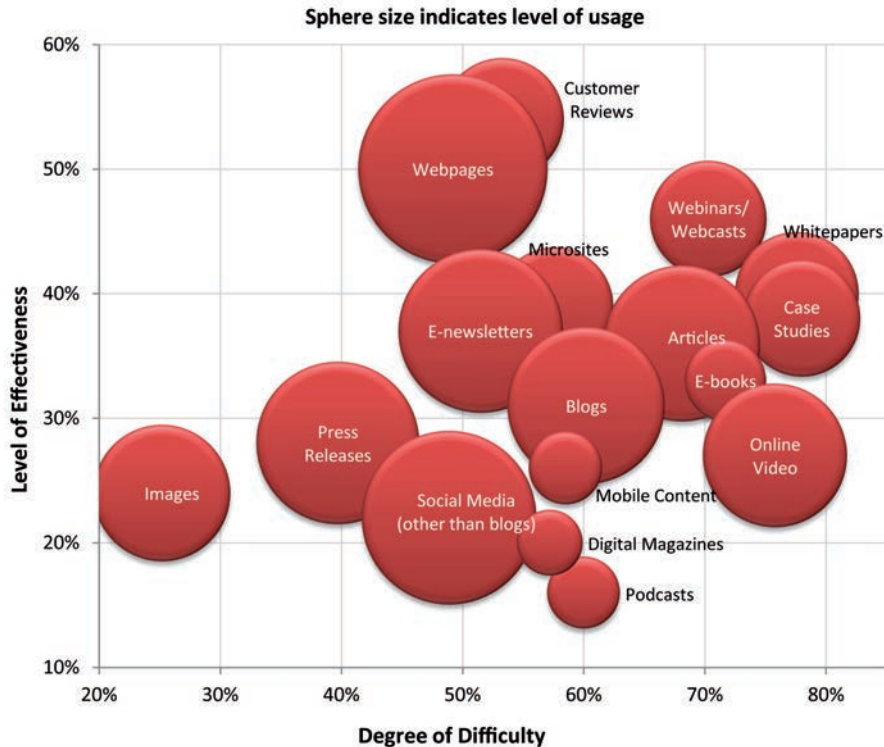
marketing**sherpa**
Source: ©2011 MarketingSherpa Search Marketing Benchmark Survey
Methodology: Fielded April 2011, N=1,530

Figure 1: Effectiveness vs. degree of difficulty of various on-line marketing channels [1].

2 Methodology

Keyword research for SEO purposes can be conducted by means of readily available on-line tools such as Google AdWords Keyword Planner [2], Bing

Keyword Research [3], ubersuggest.org, Google Trends [4], and even suggestions on the SERPs (search engine results page). These tools provide statistical information on user input (keywords) in search engine, thus, determining the most appropriate translation focused on end-users. Choosing this type of methodology, namely using on-line marketing strategies, applied to the translation process is based on the findings of several research groups that determined that the most efficient way of on-line marketing is through website content marketing (See Figure 1).

By employing such tools, translation appropriateness is determined by user usage (*vox populi*) and not by prescriptive language rules (linguistic correctness; consider misspellings, inappropriate word usage, faulty syntax, etc.) as trained in university translation courses.

Search engines reflect how vocabulary preferences shift from one period to another. Therefore, for optimal communication through the translated text, it is important to mirror the linguistic preferences of the target readers of the TT. In terms of the translation process, this step is a pre-editing process. Correctly determining during this phase the correct word base is important for the general workflow of the translation process. For instance, for the English term *website(s)*, in Romanian *site*, *website*, *sait* in the singular and *siteuri* and *saituri* are used for the plural forms, maintaining the pronunciation of the English term, whereas *sit web* and its plural *situri web* are very rarely used. By comparing the definition for the English term *site* [5] and the Romanian *sit* [6] linguists would have probably opted for *sit*, as used within the collocation *sit arheologic (archeological site).* Google Translate, probably based on statistical data, suggests *website* and *site*, whereas Bing Translator translates it as *site-ul*, adding the Romanian definite article *-ul*. In a previous study (Lakó 2009: 762-763) we showed that the preferred search term for the English *free games* was *jocuri free*. This preference faded away to the benefit of a full translation: *jocuri gratis* and *jocuri gratuite*. (Google Trends set to Romania and Romanian is useful to track user preference over time – diachronic view).

For the purpose of this paper we consider how reverse localization (Schäler 2002) can be fruitfully achieved by using the free tools mentioned above to determine the most efficient term base. On-line marketing through content marketing is based on the fact that content from web pages can be easier accessed by employing in TT words and expressions used by search engine users. Reverse localization refers to a process that is directed from a marginal language or culture (Romanian or Hungarian, etc.) to a major language/culture (English or German, etc.) We are particularly interested in Romanian to English translation and localization pre-editing processes.

3 Case Study

With the acceptance of Romania in the EU, new opportunities emerged for Romanian products and services. As a case study for this paper, we opted for "dental tourism", a booming industry in the Eastern European countries. Focus is on Romanian dental service providers that advertise themselves on the UK market, such as dental-art.ro, dentartbucharest.com, dentesse.ro with its UK URL: http://www.affordabledentistry.ca.uk, etc. However, analyzing the texts on these websites is not part of this study.

A prerequisite for a successful analysis is to set the tools to reflect information from the target market, in this case the UK market.

3.1 Open Tools for Keyword Analysis:

3.1.1 Google Adwords Keyword Planner

Google AdWords Keyword Planner (set to UK and English) is the tool to start with as it offers a reliable insight into what terms and expressions are related to the concept of dental tourism. This application provides a wide range of options to build a list of words and expressions based on a particular topic. However, using the default settings can most often offer a good insight into the keywords most frequently entered into search engines by users who are interested in such services. By default, this tool lists group ideas. The top entries are grouped under various headings and the full list contained over 800 suggestions (viewed on the 20th of August 2015).

Table 1: Partial list of suggested keywords

Dental Implants (27)
dental implants, dental implant, implants dental, how much are dental implants, dental implant procedure, dental implants uk, dental implants procedure, dental implants problems, mini dental implants, implant dental, best dental implants, all on 4 dental implants, cheapest dental implants, dental implants budapest, dental implant surgery, same day dental implants, all on four dental implants, budapest dental implants, types of dental implants, dental implant specialist…
Implants Cost (15)
dental implants cost, tooth implant cost, dental implant cost, cost of dental implants, tooth implants cost, implants dental cost, denture implants cost, dental implants costs, cost dental implants, tooth implant costs, what is the cost of dental implants, the cost of dental implants, cost for dental implants, costs of dental implants, tooth implants costs
Veneer (10)
veneers, porcelain veneers, dental veneers, veneers cost, cheap veneers, teeth veneers, tooth veneers, veneer teeth, cost of veneers, porcelain veneers cost

Dentistry (55)
cosmetic dentistry, dentistry, cosmetic dentistry prices, sedation dentistry, cosmetic dentistry cost, restorative dentistry, dentistry abroad, cosmetic dentistry abroad, implant dentistry, dentistry for you, free dentistry, laser dentistry, family dentistry, dentistry in hungary, holistic dentistry, pain free dentistry, dentistry for all, affordable cosmetic dentistry, dentistry today, general dentistry…
Teeth Whitening (6)
laser teeth whitening, teeth whitening, professional teeth whitening, zoom teeth whitening, teeth whitening dentist, cheap teeth whitening
Dentures (15)
dentures, partial dentures, dentures cost, denture, permanent dentures, denture implants, cost of dentures, dentures prices, cheap dentures, implant retained dentures, dentures in a day, affordable dentures, denture cost, cosmetic dentures cost, smile dentures
Dentist Prices (6)
dentist prices, private dentist prices, dentist price list, dentist price, dentists prices, dentist treatment prices
Cost Of Dental (24)
dental costs, dental bridge cost, dental crown cost, dental treatment costs, cost of dental treatment, dental cost, dental crowns cost, dental veneers cost, dental cleaning cost, dental treatment cost, cost of dental crown, dental care costs, dental surgery cost, cost of dental care, average dental costs, dental implant cost, cost of dental, lost cost dental care, cost dental, dental care cost…
Teeth Implants (6)
teeth implants, implants teeth, implant teeth, teeth implant, implants for teeth, implants in teeth
Tooth (18)
tooth implants, tooth implant, tooth crown, tooth whitening, tooth bonding, tooth replacement cost, tooth bridge, tooth extraction, tooth extraction cost, tooth crown cost, tooth filling, implant tooth, tooth crowns, tooth implant procedure, tooth replacement options, tooth filling cost, tooth bonding cost, implants tooth
Dental Abroad (10)
dental implants abroad, dental treatment abroad, dental work abroad, dental abroad, cheap dental treatment abroad, dental care abroad, cheap dental implants abroad, cost of dental implants abroad, dental implant abroad, dental procedures abroad
Teeth (39)
teeth whitening prices, teeth whitening cost, teeth implants cost, teeth bleaching, false teeth, teeth cleaning, teeth replacement, crowns for teeth, teeth problems, teeth crowns, crown teeth, teeth caps, teeth bonding, teeth cleaning cost, teeth treatment, cost of teeth implants, teeth inplants, teeth implants prices, crowns on teeth, teeth dentist…

Dental Practice (12)
dental practice, dental practices for sale, dental practice for sale, the dental practice, dental practices, the care dental practice, dental care practice, care dental practice, country dental practice, your dental practice, market dental practice, practice dental
Dental Tourism (30)
dental tourism europe, dental tourism turkey, dental tourism poland, dental tourism india, dental tourism budapest, dental tourism forum, croatia dental tourism, dental tourism implants, dental tourism canada, dental tourism serbia, dental tourism cuba, india dental tourism, dental tourism reviews, budapest dental tourism, dental tourism romania, dental medical tourism, vietnam dental tourism, best dental tourism, dental tourism destinations, mexican dental tourism…
Dental Plans (6)
dental plan, dental plans, dental payment plans, dental insurance plans, dental treatment planning, discount dental plans
Dental Care (18)
dental care, care dental, is dental care, emergency dental care, family dental care, dental health care, your dental care, paying for dental care, what is dental care, reasonable dental care, a-1 dental care, discount dental care, the dental care, inexpensive dental care, australian dental care, dental care for all, hungarian dental care, about dental care
Hungary Dental (9)
dental tourism hungary, hungary dental tourism, dental implants hungary, dental treatment hungary, hungary dental implants, hungary dental, dental treatment in hungary, dental hungary, dental care hungary
Dentist Cost (10)
dentist costs, dentist cost, cost of dentist, help with dentist costs, dentist costs uk, dentist implants cost, dentist treatment cost, dentist low cost, low cost dentist, dentist prices cost
Free Dental (12)
free dental care, free dental treatment, free dental, free dental work, dental treatment free, is dental care free, when is dental treatment free, is dental treatment free, dental care free, dental free, free dental near me, where can i find free dental care
Dental Prices (13)
dental prices, dental implants prices, dental price list, dental implant prices, dental treatment prices, prices for dental treatment, dental care prices, prices for dental implants, dental work prices, prices of dental implants, dental tourism prices, dental pricing, dental procedures prices
Cosmetic (10)
cosmetic dentist, cosmetic dental surgery, cosmetic dentists, dental cosmetic surgery, cosmetic teeth, cosmetic dental, cosmetic teeth surgery, dental cosmetic treatment, cosmetic dental insurance, cosmetic surgery tourism

Cheap Dental (13)
cheap dental implants, cheap dental treatment, cheap dental implant, cheap dental work, cheap dental insurance, cheap dental crowns, cheap dental care, cheap dental, cheap dental surgery, cheap dental plans, dental cheap, cheap dental clinics, cheap dental service
Dental Treatment (5)
dental treatment, dental treatments, private dental treatment, complex dental treatment, dental care treatment
Dental Insurance (12)
private dental insurance, compare dental insurance, dental health insurance, full coverage dental insurance, cheapest dental insurance, full dental insurance, how much is dental insurance, is dental insurance worth it, buy dental insurance, no dental insurance need dentist, no dental insurance, aflac dental insurance
Free Dentist (5)
free dentist, free dentist treatment, is the dentist free, free dentist care, dentist for free
Dental Clinic (8)
dental clinic, dental implant clinic, the dental clinic, dental clinics, walk in dental clinic, dental implant clinics, dental implants clinics, dental implants clinic
Dental Help (10)
help with dental costs, dental help, help with dental care, dental cost help, help with dental treatment, help with dental cost, help with dental care costs, free dental help, help for dental care, dental care help
Medical Tourism (27)
medical tourism, medical tourism uk, medical tourism thailand, thailand medical tourism, what is medical tourism, medical tourism companies, medical tourism in thailand, medical tourism statistics, medical tourism europe, medical tourism india, medical tourism definition, uk medical tourism, medical tourism poland, medical tourism agency, medical tourism destinations, india medical tourism, medical tourism providers, medical tourism dentistry, medical tourism costa rica, costa rica medical tourism…
Abroad (6)
dentist abroad, treatment abroad, dentists abroad, medical treatment abroad, medical holidays abroad, tourism abroad
Costa Rica (19)
costa rica tourism, visit costa rica, costa rica travel, costa rica adventure, travel costa rica, costa rica destinations, travel to costa rica, costa rica tourist attractions, costa rica where to go, costa rica packages, costa rica deals, tourism costa rica, costa rica trip, where to go costa rica, costa rica adventures, why go to costa rica, traveling to costa rica, implants costa rica, costa rica implants

A gist of the list shows that curating is needed. There are at least two obvious criteria to consider: relevance, on the one hand, and linguistic and

marketing effectiveness on the other. From the perspective of relevance, considering that companies under discussion are Romanian companies, keywords that contain terms such as *Budapest, Hungary, Poland, Thailand, India, Costa Rica, near me* and other non-Romanian geographical areas are not relevant. Also, keywords such as *what is medical tourism, medical tourism definition, medical tourism statistics* are clearly relevant for information only searches. All one-word keywords were also removed. This generated a list of 494 two-, three-, four-, five- and six-word keywords.
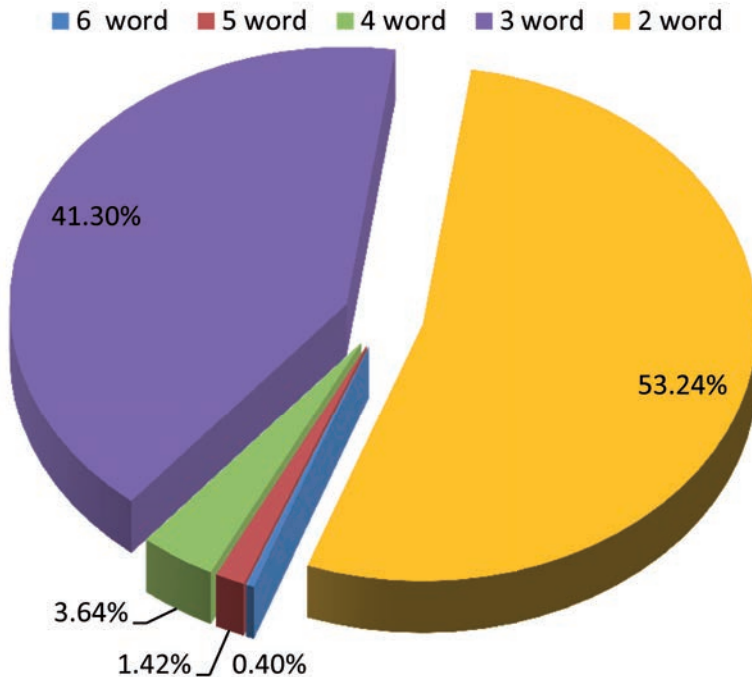
Figure 2: Percentages of keyword length suggested by Keyword Planner after initial curating from six-word to two-word keywords.

As for language usage and marketing effectiveness, several online marketing studies [7][8][9] show that long-tail keywords are more result oriented. One-, two- and three-word long keywords are not as efficient and often reflect the users' non-commitment phase. This means that users are looking for information and are only in the early stages of the buying cycle. The diagram below summarizes the views of SEO companies on the

efficiency of long-tail keywords. The longer the keyword, the higher the probability of converting a visitor into a buyer.
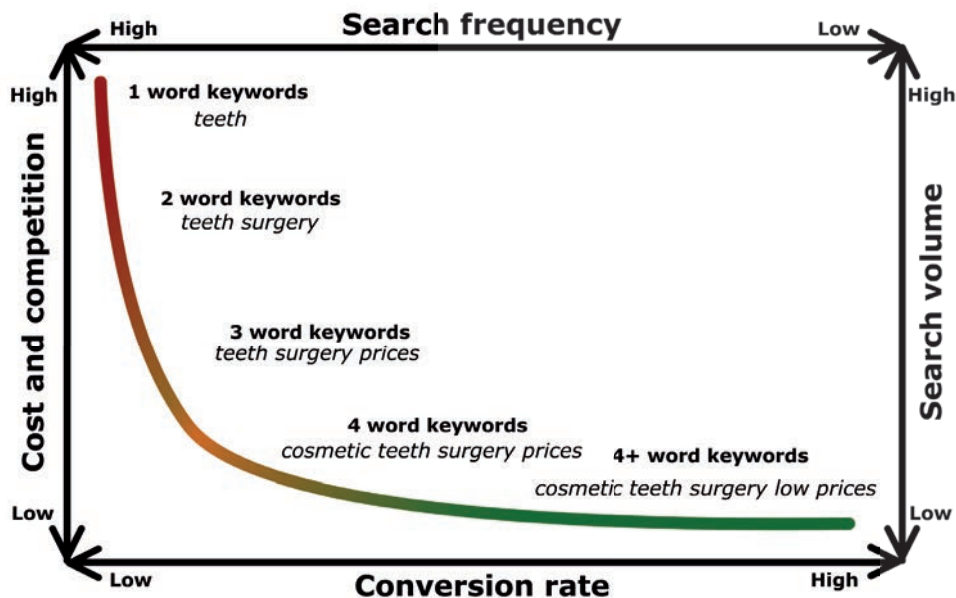
Figure 3: Efficiency of long tail keywords in web content marketing.

Considering that more than 400 suggested keywords are two- and three-word keywords, they need to be further looked up and extended to four or more words (not part of this study). This can be achieved by using various other open tools; see 3.1.2 and 3.1.3 below.

A third important factor into determining which keywords to be used in the term base is that of cost effectiveness for the potential client. For instance, *tooth/teeth whitening* procedures (using peroxide) can require lengthy periods, depending on the procedure used, and thus the beneficiary of the translation and localization can ask to remove such keywords. Probably this is why for the term *dental tourism*, a somewhat similar keyword, *tooth/teeth bleaching*, is listed only once. Seemingly, the newest whitening procedure can be effective in less than 30 minutes of treatment, during a single visit to a dentist professional. This is why it is important to check the term base against the beneficiary of the translation/localization service. Furthermore, the trans-lator/localizer can suggest terms that are rather specific to the target market, that is, the UK in this case, such as *walk in dentist, weekend dentist, dentist open on Saturday, dentist open on Sunday, dental spa, dentures in a day*.

Romanian dentist clinics may decide to implement such working strategies to come forward to the requirement of potential patients.

For marketing purposes, one can also use apparently inefficient keywords such as *affordable dental implant hungary*. The TT, as an instrumental translation process, can include phrases or subtitles such as *Romania as an affordable alternative to dental implant in Hungary,* with *alternative* as a key element in rendering the desired message, yet using a keyword very often searched for by UK search engine users.

For quick handling and quick curating Keyword Planner offers the possibility to save the suggested list as an excel file or directly to the user's Google Drive [10] account which can be used freely for curating and generating graphical data. The possible list of keywords can also be built by adding them to an advertising plan.

Also, such a list can be established by looking at the top websites that rank high in SERPS for various *dental tourism* suggested keywords. When analyzing the websites of the competitors, it is important to distinguish between the dental industry related keywords (*dental tourism, dental school, dental jobs,* etc.) and keywords that may be used by potential clients (*dental implant costs, dental implants abroad,* etc.).

Considering, for instance, *dental implant costs abroad* in google.co.uk and changing the IP (Internet protocol) address of the computer to a UK based IP (I used a free on-line IP changer [11] and accessed google.co.uk), relevant competitor web pages are displayed. Google.co.uk displays the first ten websites as if seen by a UK search engine user. Only the non-paid (organic) results should be considered (Table 2, accessed on the 28th of August 2015).

All the URLs in Table 2 can be used for benchmarking and added as an option in Google AdWords Keyword Planner to retrieve keyword suggestions that are linked to these particular web pages. As an alternative, another free useful tool from internetmarketingninjas.com [12] can be used. It can compare up to five web pages and it shows useful information such as density of one-, two-, and three-word keywords.

Moving back to the suggestions provided by Keyword Planner, the list is organized, by default, in groups. However, to remove duplicates, keywords can be sorted by keywords. For example, preference should be given to the more specific keywords (long-tail keywords). *Dental implants cost* should be listed over *dental implant*.

Considering that two- and three-word keywords are inefficient and are not cost-effective, additional tools can be employed for turning them into lengthier keywords of four or more.

1.   Affordable Dental Tourism | Patients Beyond Borders
www.patientsbeyondborders.com/procedure/**dentist**ry - În cache - Pagini similare
Sep 30, 2014 **...** Find the most trusted low **cost dental** care with Patients Beyond ... The way to achieve those goals is often an **implant**, a crown, some ... Business travelers whose work takes them **overseas** may arrange for **dental** care while on ...
2.   Dental Implants Abroad Cost £400 in Budapest - Kreativ Dental
www.kreativ**dental**clinic.co.uk/**dentist_abroad**_hungary.php - În cache - Pagini similare
Kreativ Dental offers **dental implants abroad** at a **Cost** of £400 these **cheap teeth implants**come with a lifetime guarantee at our clinc in Budapest, Hungary.
3.   Dental Implant: Affordable Treatment at Clinics Abroad - Medigo
https://www.medigo.com/en/**dentist**ry/**dental**-**implant** - În cache - Pagini similare

 Evaluare: 4,4 - 102 voturi
May 9, 2015 **...** MEDIGO lists clinics around the world offering **Dental Implant** procedures ... well connected with entire Europe by numerous low **cost** airlines.
4.   Dental treatment abroad - Live Well - NHS Choices
www.nhs.uk/Livewell/Treatment**abroad**/Pages/**Dentist**ry**abroad**.aspx - În   cache -Pagini similare
If you're considering dental treatment **abroad**, do your research and be aware of the ... For example, more than 50 different systems can be used for **dental implants**. ... by a qualified dentist before being given a treatment plan and **cost** estimate.
5.   Dentalwise - Smart Holidays Dentistry | Dentistry Abroad Clinics
www.**dental**wise.co.uk/ - În cache - Pagini similare
DentalWise is dentistry **abroad** and implant center where all your dental ... favourable rates on dental work – and **dental implants abroad** could **cost** you as little ...
6.   Dental Implants Hungary | Cosmetic Dentistry Abroad | Affordable ...
kreativ**dental**clinic.eu/ - În cache - Pagini similare
Dental Implants, Crowns, Bridges, Dentures. Quality Dental Treatment At Affordable **Prices** in Budapest, Hungary.
7.   Dentistry & Dental Implants Abroad | Treatment Abroad
www.treatment**abroad**.com/cosmetic-**dentist**ry-**abroad** - În cache - Pagini similare
Visit the Treatment **Abroad** website for guides to **dental** treatment and cosmetic dentistry**abroad** - including **costs** and quotes for treatment **overseas**.
8.   Cheap dental implants abroad - 390£ each. Lifetime guarantee ...
www.affordable**dentist**ry.co.uk/**cheap**-**dental**-**implants**-**abroad**/ - În cache - Pagini similare
Cheap dental implants abroad at 390£ each. Implants and teeth same day. **Cheap**porcelain and zirconium crowns from 190£ each. **Dental implants** with ...
9.   Dental Implants Abroad. Best Dental Implants Clinic in Bulgaria / Sofia.
dental.**implants**.bg/?visit=**implants_abroad** - În cache - Pagini similare
How much do **dental implants costs**? Check our best **prices** for **dental implants abroad**.
10. Top 10 questions about receiving dental implants abroad - Dental ...
www.**dental**departures.com/.../top-10-questions-about-receiving-**dental**-**implants**-**abroad**/ - În cache - Pagini similare
Feb 6, 2015 **...** A **dental implant** is an artificial tooth root made of titanium that is ... At a specialty clinic **abroad**, the **cost** of an implant can be as little at $1000.

Table 2: Top ten results for dental implants cost abroad, on google.co.uk (original text formatting is kept).

3.1.2 Google Search Engine Results Page (SERP)

One such tool is the Google search engine results page (SERP) itself, by entering each of the relevant two- or three-word keywords into the search field. Most Google users are already familiar with these suggestions. These suggestions show up and update as you type.

Figure 4: Google suggestions within the search engine.

3.1.3 SERP Long-tail Keywords

At the end of each SERP, Google provides related long-tail keywords.

Figure 5: Google suggestions at the end of the SERP.

3.1.4 ubersuggest.com

A useful tool that automates this task substantially is ubersuggest.com.
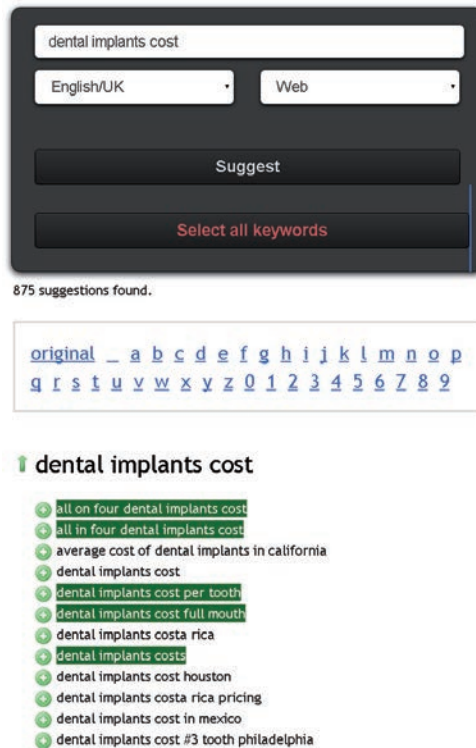
Figure 6: Ubersuggest suggestions (partial list).

For instance, if *dental implants cost* is looked up there are many suggestions that are linked to a certain geographical area, from various parts of the world that seem unlikely to be looked up from the UK, for instance *dental implants cost full mouth virginia* or *dental implants cost columbus ohio*. On the other hand, there are also quite many useful suggestions such as *dental implants cost per tooth*, *dental implants cost full mouth*.

3.1.5 Google Trends

Relevance and number of search queries and their trend can be checked and compared by using another free tool, Google Trends (set to https://www.google.co.uk/trends/?hl=en). For instance, it is important to know which the predominant keyword used should be if we compare dental implants costs vs. dental implant prices.
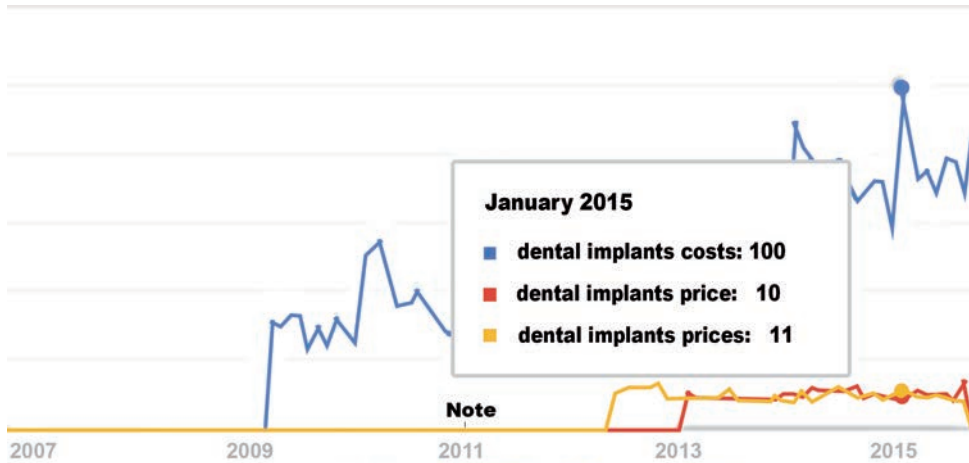
Figure 7: Comparison of various keywords as used by search engine users from the UK.

As it can be noticed, *dental implants cost* has been used ever since 2009, while the other two alternatives only later. Once all three alternatives are used, the diagram shows a clear predominance of the initial keywords. This demonstrates that some synonymous expressions should be used over their alternatives. Google Trends, as its name suggests, can also offer information on related concepts or on similar expressions. In this case, it displays the top rising keywords, reconfirming or adding to the information provided in Google Keyword Planner: *Dental implant – Medical Treatment, cost of implants, dental implants uk, nhs dental implants, teeth implants, teeth implants cost, dental implant, dental implant cost, tooth implants cost, tooth implants, dentures cost.*

3.2 Keywords as Translation Units

To a great extent, keywords found in the pre-editing stage can be considered translation units. However, the length of the translation units from the ST and the TT will not necessarily be similar. One- and two-word keywords in the ST can become long-tail keywords in the TT; moreover a two-word keyword in the ST can be efficient and cost effective since the competition in a marginal culture such as Romanian may be less fierce. On the other hand, the UK market would require long-tail keywords for successful content marketing. One impediment against associating keywords to translation units is that keywords are often unnatural sounding. Also, the on-line marketing industry considers many of the linking words that make a language sound natural as

"stop words". A list of such words can be found at:
http://www.internetmarketingninjas.com/seo-tools/seo-compare/lib/stop_words.txt

3.3 Usage of the SEO Researched Term Base

Usage of keywords in the TT should be natural, that is, in a normal way of writing. The Google indexing algorithm has evolved to such a level that it can determine if a text is overfilled with certain keywords. If the keywords are not rendered in a natural way and are meant for indexing purposes (an improper technique to fool the search engine), the web page and website is penalized. For instance, *dental implant costs romania* should be used in the TT as *… dental implant costs in Romania…*.

In order to cover as much of the potential market as possible while complying with the requirements of search engines, the translator should use predominantly the keywords that are most often used. However, synonymous expressions, related keywords, and even antonymic, yet relevant ones (see example with the keywords containing the word *hungary*), singular and plural forms should also be used. However, considering that the ST, in this case Romanian, may be very different from the TT, as the suggested approach is that of instrumental translation, rendering the text in a natural manner is of paramount importance. As the Google documentation guide suggests [13] the text should be written for the reader and not for the search engines. Due to the same instrumental translation approach TT text length will vary from that of the ST. Also, in terms there is a good policy to check the text length particular for a certain web market segment in the target culture.

3.4 Rage Against the Machine in Translation

The term base built using the open tools described above can be used in translation memories (TM) for automating translations. However, in the case of web content marketing, using and overusing the same keywords (even more so if we consider the long-tail keywords) can result in penalization from search engines. Using Wikipedia or other free community-driven websites for building a term base for a specific field of human activity can also lead to non-voluntary plagiarism. This can occur from overusing such sources that make up a translation memory. In order to be indexed in search engines, it is important that the content be new and original in the target language.

Also, in theory, articles may require "rewriting" by using new predominant keywords, or adding alternatives (see Google Trends); however, the life cycle of articles is usually shorter than the life cycle of certain keywords (*dental implants cost* vs. *dental implants prices*). As a counterexample, keywords that contain a time stamp have a reduced life cycle and so do the articles that

contain them; consider *dental implant costs 2015*. While it reflects updated information, its life cycle is limited to 2015. Search engines value unique, updated, and valuable content, so there is not much room for automatisms.

4 Conclusions

This type of approach to the pre-editing translation process is beneficial as it provides reliable statistical data, and can be applied successfully especially to web content marketing. The tools needed to achieve such translations are free to use and therefore can be used by anyone, from freelancers and small companies to multinationals. For determining the most lucrative set of keywords, moving back and forth with each of these tools may be required.

By employing a marketing approach to instrumental translation, the beneficiary of the text gains a competitive edge over its competitors; hence, the outcome is a value added translation. Pym (cited in Dimitriu 2002: 98), suggests moving from a purely linguistic perspective to a sociological and economic one, as in the case of websites, more often than not, the driving engine is generating sales. Building texts based on the language expressions used by the potential clients opens up more efficient communication channels. Also, this approach implies a rather copy-writing-like process, namely moving further away from the ST. The main benefit is that the TT is far less under the influence of the ST which makes integration into the target culture much smoother.

Regarding the applicability of this method, for the purpose of this paper we considered Romanian as the source language/culture and British English as the target language/culture. However, this method is reusable and reproducible with any language/culture pairs and can be applied to any industry by using the same open tools or similar ones.

References

Dimitriu, R. (2002) Theories and Practice of Translation. Iași: Institutul European.

Dimitriu, R. (2009) Key words and concepts in E. A. Nida's approach to translation and their further development in Translation Studies. In Dimitriu, R. and Shlesinger, M. (eds.) Translators and Their Reader, Brussells: Les Éditions du Hazard, 23-41.

Lakó, C. (2009) Translating for the Web – The Keyword Oriented Translation Process. In The Proceedings of the European Integration – Between Tradition and Modernity Congress (3), Târgu-Mureş: Editura Universităţii "Petru Maior", 761-767.

Nord, C. (1997) Translating as a Purposeful Activity. Functionalist Approaches Explained. Manchester, UK: St. Jerome Publishing.

Nord, C. (2005) Text Analysis in Translation: Theory, Methodology, and Didactic Application of a Model for Translation-Oriented Text Analysis, Amsterdam: Rodopi.

Schäler, R. (2002) Reverse Localization. The International Journal of Localisation, Vol.6, Issue 1, Localisation Research Centre, CSIS Dept, University of Limerick, Limerick Ireland, 39-48.

Internet sources:

[1]https://www.marketingsherpa.com/data/public/reports/special-reports/SR-A-Tactical-Approach-to-Content-Marketing.pdf

[2] https://adwords.google.com/KeywordPlanner

[3] http://www.bing.com/toolbox/keywords

[4] https://www.google.com/trends/

[5] http://www.merriam-webster.com/dictionary/site]

[6] https://dexonline.ro/ definitie/sit

[7] http://ds6.net/wp-content/uploads/2014/05/LongtailEbook.pdf

[8] http://neilpatel.com/2015/05/07/a-step-by-step-guide-to-integrating-long-tail-keywords-within-blog-posts/

[9] http://www.business2community.com/seo/secrets-using-long-tail-keywords-01154759#bZAmD3APR239xmfI.97

[10] https://drive.google.com

[11] http://www.onlineipchanger.com/

[12] https://www.internetmarketingninjas.com/seo-tools/seo-compare/

[13] http://static.googleusercontent.com/media/www.google.com/ro//webmasters/docs/search-engine -optimization-starter-guide.pdf

Openness in Computing
The Case of Linux for Translators

Peter Sandrini
University of Innsbruck, Austria

The decision to use exclusively open source software for translation purposes includes deploying an open source operating system. Put in another way, if I, as a translator, want to use free and open source applications on my PC, it is legitimate and almost obvious for me to support this choice by using an open source operating system as well. An operating system constitutes the basic infrastructure of any computer system: without it, no application can be launched and no data can be edited or saved.

In this context, openness first and foremost means using a free and open source operating system, thus, eliminating the need for any proprietary software; secondly, openness is also about having the opportunity to be part of a community, by sharing and contributing one's own experiences and solutions.

The following paper describes the use of GNU/Linux as a platform for translation, summarizes experiences and opportunities, and gives a historical overview over different initiatives trying to adapt the GNU/Linux environment for translation.

1 GNU/Linux – The Operating System

GNU/Linux is a piece of software "that enables applications and the computer operator to access the devices on the computer to perform desired functions" (Linux Foundation 2015). It represents the deep software layer of a computer systems on which all other applications build upon. What sets GNU/Linux apart from comparable commercial solutions, such as Microsoft's Windows or Apple's OS X, is the collaborative development based on a community of programmers who contribute to the system. Nobody owns GNU/Linux and there is no single company responsible for GNU/Linux even though a few commercial companies contribute code on a regular basis; there are, however, numerous communities, each working on a specific component of the system.

The story of GNU/Linux begins when in the late 1970s a programmer at MIT, Richard Stallman, became dissatisfied with the increasing commercialization of the old UNIX computer operating environment. He began to develop a set of tools, called the GNU (GNU Is Not Unix) tools, as a first step

on the way to a free operating system. While the main tool-set was ready rather quickly, the central part of the operating system, called its kernel, was still missing and the corresponding HURD kernel project lagging behind time. In 1991 the Finn programmer Linus Torvalds programmed a new kernel and gave it the name Linux. Thus, the Linux kernel successfully complemented the GNU tools and became the core architecture of a complete and open source operating system, the GNU/Linux system (Stallman 2014).

There are general arguments in favor of GNU/Linux over other OSs: over the years, it has become a stable and mature operating system which can easily replace any other system. A strong emphasis on security, for example, makes anti-virus software more or less obsolete, a robust system architecture avoids frequent rebooting, thus increasing efficiency and productivity.

These general advantages, however, may not be the main reason for a change to GNU/Linux; it is its openness and free availability, giving the user a choice of more than 500 different flavors of Linux distributions. GNU/Linux relies on the work of communities, it is free software and as such it is subject to the four essential freedoms as defined by the Free Software Foundation (outlined in the introduction to this volume). With these freedoms, the users, both individually and collectively, gain control over their computers and the technology they use:

- Users can be assured that their computing remains confidential as the code is open and back-door attacks to the system are immediately detected and removed.
- The integrity of the program code is guaranteed through its openness.
- The integrity of user data is guaranteed through the stable system architecture and the almost complete absence of viruses.
- Users have complete freedom over installation and configuration of software.
- Users have a choice and can be part of a community, changing from dependent consumers of a purchased product into active and autonomous agents, completely independent of commercial interests and big companies.

The advantages of having full control over one's own PC includes ease of computer installations without having to input activation codes or managing software licenses. Still, there is no fear of copyright infringements even when multiple instances of the system are installed, e.g. on a desktop and a notebook computer, or in a computer lab of a school or university. For students and university graduates full control also allows a cost-saving start of

their professional career which is especially important during a first orientation period.

Openness and control over the computer system also facilitates co operation with colleagues by eliminating the risk of malware and viruses, by supporting open standards, as well as by fostering discussion and exchange through participation in on-line communities in support of free and open source projects.

1.1 Language Support

In addition to having control over their own computer, users may count on a rather extensive language support, in many cases exceeding that of commercial operating systems. The mainstream GNU/Linux distribution Ubuntu, for example, supports around 150 languages: it comes in English by default, but users may choose from more than 146 additional languages to install, and get the user interface in their mother tongue. This originates from the fact that GNU/Linux developers are organized in many individual projects scattered all over the world, so that language support even for smaller and less developed locales was recognized as a necessity right from the beginning. For this purpose, a thorough localization method has been introduced for the operating system as well as for all applications meant to run on it: the GNU GETTEXT environment, designed to minimize the impact of internationalization and localization on the program source code.

Specifically, the GNU GETTEXT utilities are a set of tools that provide a framework within which free software packages can produce multilingual messages, as well as a set of conventions about how programs should be written to support message catalogs. These message catalogs, called PO files, contain both the English and the translated versions of each message. PO stands for Portable Object, distinguishing it from MO files or Machine Object files. PO files are meant to be read and edited by humans, and associate each original, translatable string of a given package with its translation in a particular target language. PO files are strictly bilingual, as each file is dedicated to a specific target language. If an application supports more than one language, there is one such PO file per language supported.

The utility program XGETTEXT creates a PO Template file (POT) by extracting all marked messages from the program code sources, the MSGINIT tool converts it into a human readable PO file. Another utility, MSGMERGE, takes care of adjusting PO files between releases of the corresponding sources, excluding obsolete entries, initializing new ones, and updating all source line references. Translators then edit and translate the

messages contained in the files with the help of simple text editors or dedicated PO file editors such as Lokalize, the PO file editor of the KDE desktop environment, Gtranslator and PO-Edit from the GNOME desktop environment, or PO Mode, a specific add-on for the text editor Emacs. PO files are only used as an intermediate file format in the development and localization process: after translation, the MSGFMT tool converts PO files to binary resource files, or MO files, which are then used by the GETTEXT library at run time.

The GNU GETTEXT environment was one of the first thorough software localization methods and it was introduced by the free software community and the GNU/Linux system in 1995. PO files also constituted the first translation data format long before XML formats such as TMX and XLIFF were invented. The localization of free and open source programs is well supported and documented; the excellent introduction written as a Master thesis by Arjona Reina (2012) explains the process in detail and gives an overview over tools and platforms.

Once translations are in place, users can influence the language used by the operating system and by installed applications in different ways:

1. During the installation of the system, users may choose a preferred language which sets the system-wide default language for all users, as well as the language used when a new user account is created; each user can have his own locale configuration that is different from the locales of the other users on the same machine.
2. By setting the GUI language of a desktop environment, such as KDE, GNOME, or XFCE, which usually includes the window manager, a web browser, a text editor, and other applications. The locale used by GUI programs of the desktop environment can be specified in a special configuration screen.
3. By configuring a series of environment variables like LANGUAGE, LC_ALL, LC_XXX, LANG.

In addition, text input can be adapted to different writing systems by installing specific tools and setting up the operating system accordingly. Furthermore, Unicode, the Universal Character Set standard, capable of encoding, representing, and handling of text expressed in most of the world's writing systems, has become standard in most GNU/Linux installations.

Because of the GNU GETTEXT environment and the versatility of configuration options, modern GNU/Linux distributions are indeed well suited as multilingual computer systems for everybody who needs to use, write or work with two or more languages.

1.2 Adoption

Today, most users who face a GNU/Linux system for the first time already had some experiences with a proprietary operating system. A change of the main operating system involves a certain degree of readjustment: new interface, new commands, new system applications and a new way of organization have to be learned. The whole change may be represented as "trading Windows problems for Linux challenges" (Hartley 2015). GNU/Linux is not more difficult to handle than other proprietary systems (see survey results in García González 2013: 141) as it has often been blamed, it is just different, and users have to adjust. This initial difficulty is often mistaken for greater complexity, but it is not, as GNU/Linux users who return to using a proprietary system, very often encounter the same challenge.

GNU/Linux comes in a variety of distributions, each one with its particular features, some even geared to a specific task. The main distinction to be made, however, is the discrimination of three specific areas of use: as a server operating system, a desktop system or a mobile operating system. While GNU/Linux on servers has a share of 36% for public servers on the Internet and 97% for supercomputers in 2015 according to Gartner research (Wikipedia n.d.), and Linux on mobile devices, including Android which uses the Linux kernel, tops all other operating systems, it struggles to achieve the same results on the desktop. Adoption rates on desktop systems are very hard to get and in most cases the operating system is identified by web counters. The figures coming from such web counters attribute a rather small market share to GNU/Linux: from 1.47% for 2015 (Net Market Share n.d.) to around 5% (W3schools n.d.). The Linux Counter Project website (Linux Counter Project n.d.) describes the difficulties in assessing exact numbers of users, but estimates the number of GNU/Linux users worldwide at 79,879,362.

The number of users of specialized GNU/Linux distributions, such as the distributions for translators mentioned below, are even harder to assess: there are numbers of downloads, e.g. from Mediafire where *tuxtrans* is hosted, or the number of participants and messages in on-line discussion groups, but they all only indicate trends, show interest, but they do not give evidence of the number of actual users. In view of available numbers, even if these data is highly unreliable, we have to conclude that, basically, GNU/Linux remains a niche operating system on the desktop, and, thus, also in the translators community.

However, several initiatives and projects have recognized the advantages and usefulness of free and open source software in general, and on the desktop in particular. The European Union's Open Source Software Strategy

2014-2017, for example, states that the EU "Commission shall continue to adopt formally, through the Product Management procedure, the use of OSS technologies and products", in order to "ensure a level playing field for open source software and demonstrate an active and fair consideration of using open source software" (EU Commission 2015: art 1 and 2). For this purpose, several initiatives were launched within the EU, e. g. the Joinup collaborative platform (EU Commission 2015b) aiming at interoperability solutions for public administrations, formerly called OSOR, the Open Source Observatory.

2 GNU/Linux for Translators

When we speak of a free operating system for translation and translators, we need to specify this particular target group more clearly. Translators may be single free-lance translators, working on a desktop computer, they may be translating voluntarily in their spare time for non-governmental organizations, open source software or charity projects, or they may work for a translation agency as professional translators to earn a living. In today's globalized world, all translators rely on some form of networking or Internet use, be it the exchange of translation memories or other language data between voluntary translators, the use of on-line or cloud-based translation tools, such as Google Translator Toolkit (GTK n.d.), Dotsub, a cloud-based subtitling platform (Dotsub n.d.), the Trommons, a web-based translation environment developed by The Rosetta Foundation (The Rosetta Foundation n.d.), or even the use of on-line translation memories tools such as Matecat (Matecat n.d.), Linguee (Linguee n.d.) or MyMemory (MyMemory n.d.) or on-line term banks. Networking and Internet use, however, become a necessity for professional translators for which cooperation and on-line presence is a must: Cronin speaks of the network-based nature of the translation industry "where translation projects are managed across countries, continents, cultures and languages" (Cronin 2003: 45).

Translators are, thus, a diverse target group and translation is far from a homogeneous activity. Yet, some common features and prerequisites for a computer system suitable to the task of multilingual communication and translation may be identified:

- Wide multilingual support – language support comprises a wide choice of languages for the user interface of the OS, language support for installed applications, support for different text input systems and language-specific keyboards layouts.

- Support for standards, especially standards regarding multilingual text (Unicode n.d.), writing systems (text input, fonts), translation (PO, TMX, XLIFF), terminology (TBX).
- Inclusion of translation technology applications: CAT, MT, TM, Terminology, etc.

For a translation-oriented computer system, specific applications must be included, configured and installed. The operating system only represents a platform for applications of translation technology, and translation technology applications run on such a platform. This creates a mutual dependency: an operating system for which no specific translations-oriented applications exist is of no use, and software applications will not work without the basic infrastructure of the operating system.

Technology has become indispensable in many areas, for translation scholars even speak of a "technological turn" (Cronin 2010: 6). Translation technology has been around more or less thirty years now, but the number of available software products as well as specific free and open source projects has multiplied in recent years. Translation technology may be defined as any kind of digital Information and Communication Technology (ICT) which supports or performs the translation process with the aim of meeting adequate efficiency and quality requirements (Sandrini 2012: 111). While for many translators, translation technology still equals Trados, a widely used proprietary CAT tool, a number of specific translation-oriented applications have been developed exclusively for or ported to the Linux environment, so that today there is a variety of options available. Commercial products, such as Swordfish, Wordfast Pro, Cafetran, XTM, MemSource and others, are available on the market on the basis of proprietary licenses, and more importantly, a plethora of free and open source software applications is listed in the FOSS4Trans catalog (FOSS4Trans n.d.) with no less than 150 specific programs for GNU/Linux subdivided into four broad categories:

37 editing and publishing tools (plain text and code editors, office suites, desktop publishing, advanced image editors, subtitle editors, optical character recognition software, differencing tools, PDF tools);

30 language tools (terminology extraction, text analysis, corpus creation and processing, resource lookup tools, language checkers);

59 translation tools (translation environments, machine translation programs, localization tools, alignment tools, format conversion and validation utilities for translation-related formats);

24 management tools (project management programs, word-counting and invoicing tools, financial management software, reference management tools, quality assurance tools).

Although some items in the list do not strictly qualify as translation tools proper, and some software projects have already stopped development, this catalog is good evidence of the availability of translation technology applications for the GNU/Linux platform. Thus, the argument often brought forward against GNU/Linux that there are too few CAT tools for this platform is no longer valid.

There is, however, one main difference between some proprietary systems and GNU/Linux, or more in general, Unix. Applications developed for this platform mainly follow a specific design principle which goes "make each program do one thing well" (McIlroy et al 1978: 1902), as outlined early by its main programmers. Programs are designed to do one thing and do it well, so that applications focus on just one specific task. As a consequence, in the free and open source community we have a great number of individual projects creating applications with specific functionality on the basis of this principle: one or more communities developing a terminology management system, others developing terminology extraction tools or terminological/lexicographical file format converters, projects dedicated to spell checking routines, communities developing text format conversion tools, others creating translation project management programs, yet others implementing accounting software for translators, and so on. This splitting-up of human resources is somewhat attenuated by a second GNU/Linux and Unix programming principle which says: "write programs to work together" (McIlroy et al 1978: 1902), making the output of every program the potential input to another program. Communication and data exchange between programs, thus, becomes crucial and of central importance. So, you may end up with many different tools but they are all able to interconnect in one way or another.

Opposed to a great number of translation technology applications in the free and open source world each of which concentrating on one specific functionality, we find huge all-encompassing software programs, called Translation Environment Tools (TEnT) (Zetzsche 2014: 189) in the proprietary and commercial world. Such a computer-aided-translation tool or TEnT aims at providing a one-stop solution for translators with all needed functionality, from a translation-memory engine, terminology management, alignment, collocation search and translation project management, up to format conversion, spell-checking, text editing and formating tools, etc. This results in having only a few contenders for market leadership in the commercial environment, but a great number of projects and communities in the free and open source world. Translators exploring free and open source programs should get accustomed to the thought that there is more than one program for a specific task and that they are expected to try out and combine different applications for a useful translation workflow.

Translation Technology can boost the efficiency and consistency of translation, but inconsiderate use of software and services may also cause translators losing control over the translation process and translation data. This is especially true for web-based translation tools, software-as-a-service applications (SAAS) and closed source programs. Costs and risks of software technology should always be evaluated: this could give free and open source tools a clear advantage due to their very low access barrier with regard to costs, and their security and reliability.

In combination with the principle of openness, the specific features and advantages of the GNU/Linux operating system may be summed up as follows:

- a stable and secure operating system
- good usability, easy handling and configuration
- full control over the system
- variety of available FOSS applications
- no financial costs, free to use and free to redistribute.

A change to GNU/Linux, however, always involves a certain degree of rethinking one's habits and practices using the computer, it demands adaptation, and could be a learning challenge. On the other side, such a change opens up a new world of unfettered use of the computer and, according to the Distrowatch.com website, it puts "the fun back into computing" (Distrowatch.com n.d.).

2.1 Linux Distributions for Translation

Acknowledging these advantages, a few initiatives promoting the use of GNU/Linux in translation were launched in the last two decades. One of the first initiatives was a website (Prior 2010) created by Marc Prior around the year 2000 in which he reports on his experiences using GNU/Linux as a translator in his day-to-day work. He lists and describes applications of inter-est to translators, shares his experiences and offers links to many GNU/Linux-related websites. Hand in hand with this website a discussion group was created on Yahoo, the Linux for Translators Forum, "intended primarily for pro-fessional translators who use GNU/Linux software for their work" (The Linux for Translators Forum 2002: About Group) with 614 members at the time of writing. Discussions in this group address all topics regarding the use of GNU/Linux for professional translation tasks.

Linux for Translators Forum: n° of messages
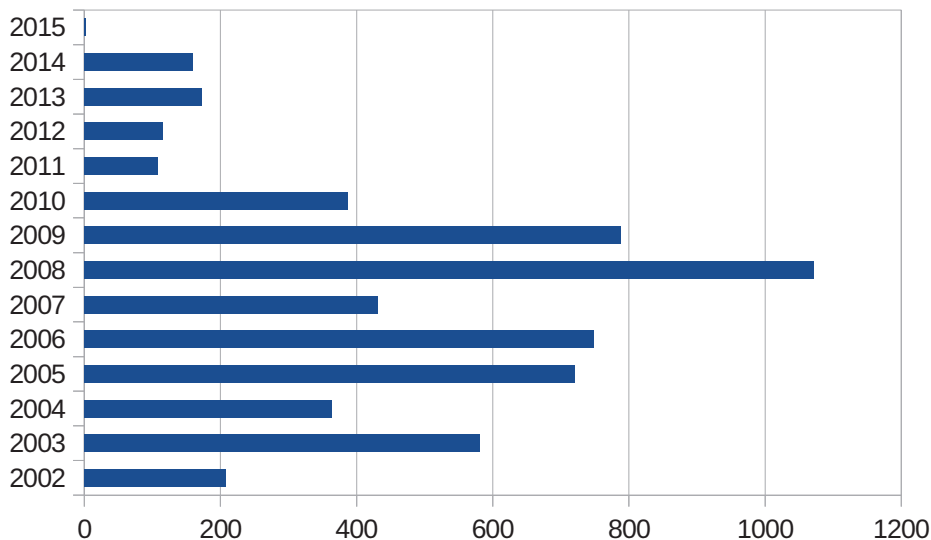
Figure 1: Number of messages in the Linux for Translators Forum.

Figure 1 shows that interest in this forum peaked in 2008 with more than 1000 messages, and settled down to around 150 messages a year after 2011. The time around 2007 and 2008 also was the beginning of dedicated GNU/Linux distributions for translators while the first years of the millennium saw the development of some of the most prominent free and open source applications in translation like OmegaT, Pootle, Open Language Tools, Apertium, Moses, Globalsight and others.

The use of GNU/Linux has also been the topic of discussions in user forums or websites dedicated to professional translation such as ProZ.com and Translatorscafe.com. Starting from 2005, the "Festival Latinoamericano de Instalación de Software Libre" (FLIsoL n.d.), a series of regular events in Latin America, promotes the use of free software and free culture, organizing among other things workshops about the use of Ubuntu and *tuxtrans* for students.

In 2007, the group GETLT (Grupo de Estudos das Tecnoloxías Libres da Tradución (GETLT n.d.) was created at the University of Vigo, Spain, with the following goals in mind: to analyze and promote the use of free software in professional translation practice, as well as in translator training; promote the

visibility of the work done by volunteer translators of free projects; stimulating the cooperation of students, teachers and professionals in translation with communities involved in translating free software projects. In addition to relevant publications (Diaz Fouces et al. 2008, Diaz Fouces 2010, García González 2013), the main product of this group was the development of a translation-oriented GNU/Linux distribution called MinTrad.

Distributions are software packages which include GNU/Linux, as well as a number of selected applications. A GNU/Linux system, with its set of tools surrounding the kernel, different window managers and a great number of complete desktop environments providing the graphical user interface (GUI) and allowing the interaction with the user, is very modular, and for each component, numerous projects have developed slightly or totally different compatible versions which may be exchanged at the discretion of the user. Due to this modularity, a GNU/Linux system may be configured and set up in many different ways, for different tasks and different environments. This generated several independent releases of GNU/Linux called distributions – Distrowatch.com lists more than 500 of them – where the distribution's makers which may be a company, an individual or a community have decided which kernel, operating system tools, environments, and applications to include and ship to users.

A few attempts have been made to tailor a GNU/Linux system to the requirements of a translator, making choices with regard to two different aspects: 1) decision about which system tools, window managers or desktop environments to include, and 2) decision about what applications to configure and install. Ideally, both decisions should be based upon how well multilingual support, open standards and translation technology are supported; but in some cases, e. g. the choice of a desktop environment like KDE, Gnome or XFCE, it may be a matter of personal preferences.

The following GNU/Linux distributions have been developed explicitly for translators and include free and open source standard applications like web browsers, email clients, office suits – mostly LibreOffice –, as well as dedicated translation technology software, such as translation memory systems – mostly OmegaT –, terminology applications and text analysis programs. These categories are the most commonly used applications by professional translators (see survey results in García González 2013: 137).

LinguasOS

LinguasOS was developed by Tony Baldwin, a "translator and translation agency owner who is intimately familiar with the needs of professionals in the translation trade" (Baldwin 2008) in December 2007. It is a based on

PCLinuxOS, more specifically on PCFluxboxOS with the minimalistic window manager Fluxbox, and adapted for professional translators and those working in software localization with many specific applications and support for all industry standard file formats.

LinguasOS "a) attempts to give translators a platform for experimenting with the tools that are available in FOSS for the trade, in a quick and light Live CD distribution, as well as, b) provides an easily maintained, preconfigured OS for translators that are already using, or wish to begin using Linux for their work" (Baldwin 2008).
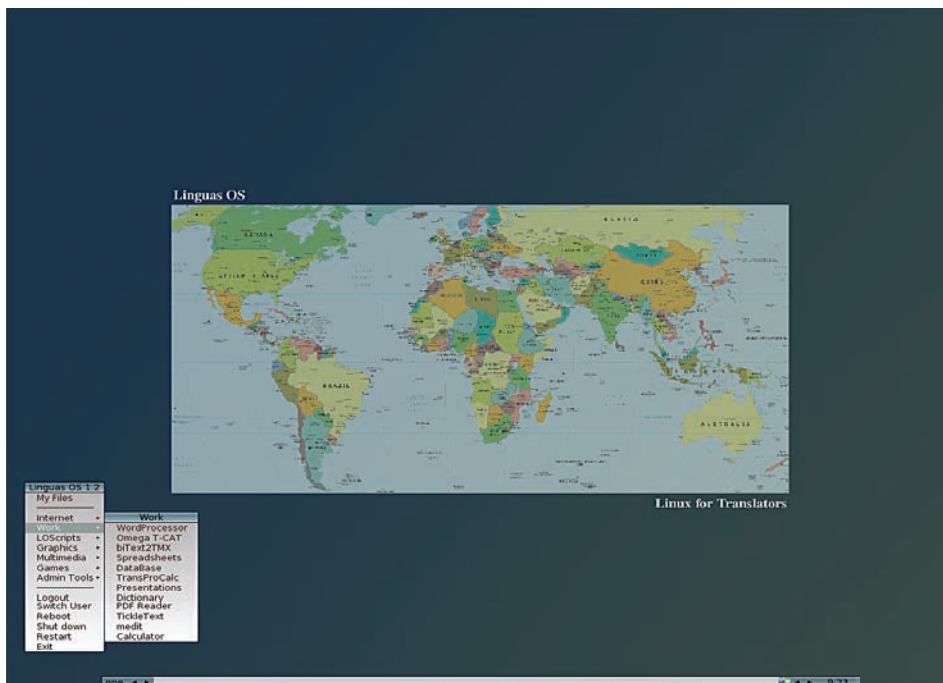
Figure 2: LinguasOS start screen and application menu.

The system comes as a live-CD packaged in an ISO file with only 412 MB of disk space which can be started for trial purposes from a CD or a USB stick without installing or changing anything on the computer; however, installation on the hard-disk is also possible.

The user forum (LinguasOS discussion group n.d.) has messages going from December 2007 through February 2010. LinguasOS is still listed on

Distrowatch.com with its most recent release (1.3) dated March 2008, though development was officially stopped in October 2009.

MinTrad

From 2007 to 2010 the GETLT group at the University of Vigo in Spain launched a project with the title "Creation of a GNU/Linux training environment for the training of translators, software localizers and subtitle editors" (see García González 2013: 130 and Veiga Díaz/García González in this volume) financed by the Galician Regional Government, with a slightly different target group focusing on academic teaching, and widening the con-cept of translators to include multilingual communication and localization. The resulting distribution MinTrad is based on Linux Mint and features a traditional desktop with a custom menu item 'MinTrad' listing all translation-specific programs.

   The Linux Mint basis represents a user-friendly and reliable system and the choice of programs is well thought-out, even though OmegaT comes in three slightly different versions (OmegaT, Autshumato, OmegaT+). However, the last version in the download section of the FTP-server (ftp.uvigo.es/mintrad/) accessed at the time of writing, is dated September 2012.
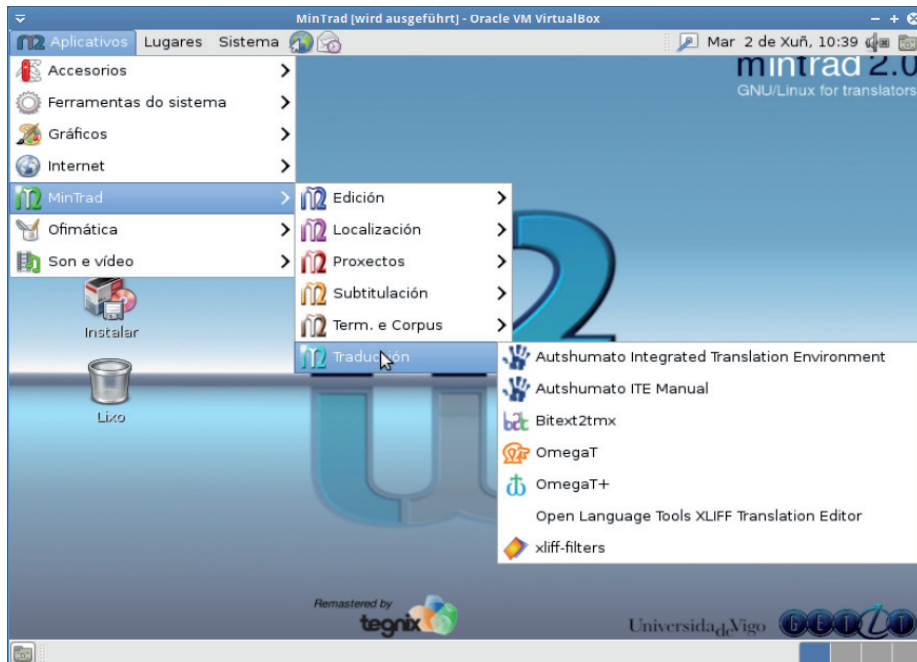
Figure 3: MinTrad start screen and application menu.

tuxtrans

More or less at the same time, in December 2007, another customized system for translators called PCLOSTrans was created at the University of Innsbruck in Austria. It was based on PCLinuxOS a more general GNU/Linux distribution featuring the KDE desktop; in 2010 this basis was exchanged for the widely used Ubuntu distribution with both the XFCE and the Fluxbox desktop and the name was changed to *tuxtrans*. The most important open source applications of translation technology are included and made accessible through the customized menu 'Translation'. The user interface is available in four languages English, Italian, German and Spanish, but more may be installed on-line from the Ubuntu repositories.

A dedicated user forum (*tuxtrans* discussion group n.d.) lists messages going from May 2010 through January 2015, and the tuxtrans website (tuxtrans n.d.) has introductory notes on how to install and use the system, as well as a FAQ page. The system comes in a 32bit and a 64bit version and the last update available for download is dated September 2014 (32bit) and March 2015 (64bit).
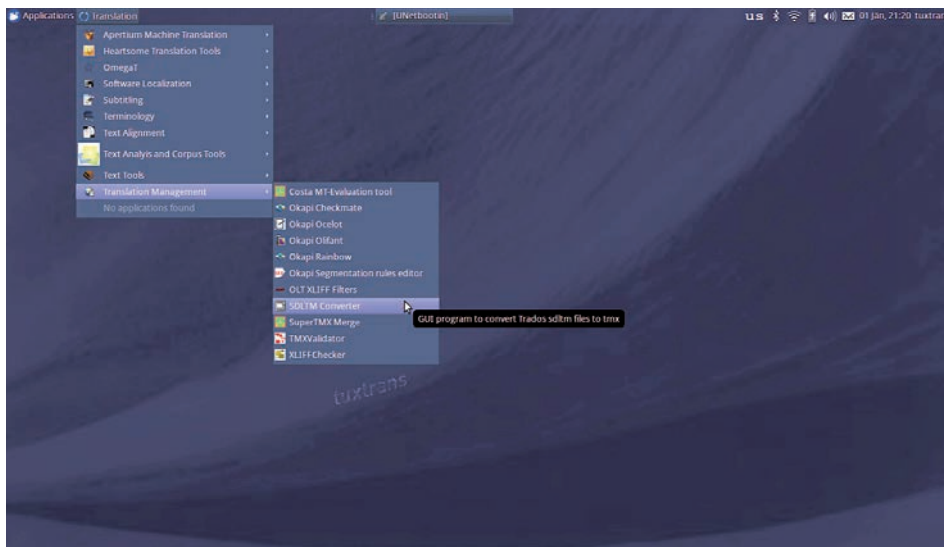
Figure 4: tuxtrans start screen and application menu.

Apart from standard applications and the most common translation technology programs, the three distributions differ in their integration of machine translation and locally installed web-based applications. Every

GNU/Linux system can act as a Web server if the right software is installed. Thus, it would be an interesting future path to move from a single-user desktop system to a distribution which already includes the necessary infrastructure software (e. g. databases, web servers) coupled with multi-user translation technology applications, such as for example, the translation management system Globalsight, the terminology management system Autshumato TMS, a multilingual web content management system like Drupal, a translation server like Pootle, etc. Such a GNU/Linux system can be used either as desktop system or, when properly installed, as a multi-user translation server, or with the appropriate hardware, even both uses at the same time on the same machine are possible.

For machine translation, there is already Apertium working off-line which can be installed very easily as a Java application with all the language combinations supported. The Moses MT system, another open source machine translation system, requires much more effort and know-how for installation, and, in particular, plenty of disk space for a working instance with one language combination, and each new language combination adds further disk space; installation of such a language-specific, or better language-combination-specific program in a general, translation-oriented distribution, therefore, does not make much sense without a limitation to two working languages.

3 Conclusion

Even though two of the three distributions are not updated any longer, these projects still prove that using exclusively free and open source software does constitute a real option for any kind of translator, allowing her to do all relevant tasks in translation and localization. Nonetheless, there is still no sign of a wide adoption of GNU/Linux as an operating system for translators, and no major breakthrough has been made, at least judging from GNU/Linux adoption in general, and direct feedback, questions and reports from users of *tuxtrans*, in particular.

With all the advantages mentioned earlier, the robustness of the system, the possibility of easy testing with live-systems booting from a DVD or a USB stick, and, last but not least, the negligible cost, the question has to be raised what factors prevent users in translation, localization or multilingual communication from adopting GNU/Linux. A few possible reasons can be tentatively mentioned:

- Reluctance to change to a new operating system from the old accustomed one; in many cases, new computers come with a pre-installed proprietary operating system, and in many companies or

institutions only one proprietary system is supported, so that users usually start out with this system and get used to it, thus increasing the barriers for a change;
- Assumed or real complexity of GNU/Linux;
- Absence of professional support; with GNU/Linux there is nobody to blame when something goes wrong, except one's own knowledge and preparation. For some users the change from commercial support to voluntary support through communities may pose a challenge;
- Incompatibility of specialized software; not all software programs run under GNU/Linux. That being said, the best approach would be to look for comparable (functionality) and compatible (support for standards) applications, i. e. when a user says "I cannot use GNU/Linux because Trados does not run on it", the right question to ask would be "What are the reasons for using Trados?" and "Could the free translation memory system OmegaT replace it?" as well as "Can you exchange translation memory files on the basis of TMX?" In many cases this could solve the problem, provided there are no other more serious reasons;
- Lack of awareness and not knowing about GNU/Linux and free and open source options: this is, among other things, why this article is written. Poor knowledge about GNU/Linux and free and open source software in general among translators and translation students has been mentioned in a survey conducted in 2008 (García González 2013): "the almost complete unawareness of the characteristics and possibilities of open-source software revealed by the participants" (García González 2013: 136).

All this reasons could deter translators from using GNU/Linux, but it is nearly impossible to identify the most important, or the most influential factors. Instead of guessing what keeps users/translators from using GNU/Linux, maybe everyone, or every computer user should better ask: Why should I not use a free and open source operating system that is freely available, secure, multilingual and ready to be used for translation? And: Why should I, then, pay for a proprietary operating system?

With no clear picture about the key factors influencing user adoption, it could be useful to identify common measures that address all of these factors. Intervening on the last one, i. e. to inform and educate potential users about free and open source software and operating systems, seems to be also at the heart of the other reasons where a lack of knowledge or understanding constitutes a major problem. This is done primarily through public promotions by non-governmental organizations, like the Free Software Foundation (FSF

n.d.), the Free Software Foundation Europe (FSFE n.d.), and others, or by personal initiatives.

Promoting and enhancing information and awareness about free and open source software could be done best within academic organizations and translator training institutions where future translators learn about the choices and options they have when it comes to translation technology and basic IT infrastructure. Narrowing down their options to proprietary systems would not be in accordance with good academic practice: teaching at university level is, indeed, more about empowerment of students than simple product training (Diaz Fouces 2011). All the advantages of using free software in education (FSFE n.d.) apply to the use of a free and open source operating system as well: no license fees, no trouble with licenses, equality for all students and teachers, etc, in addition to the general advantages mentioned earlier in this contribution.

With this in mind, we may answer the question why the makers of specific distribution of GNU/Linux for translators are doing this work and providing such a system for free. From personal experience, I would say, they do it as prove of practicability, because it can be done, or even because somebody needs it. In the case of *tuxtrans*, the fact that it actually represents the system I am working with myself, greatly facilitates the production of this distribution. Independently of the number of potential users, free software allows me to make my desktop computer – operating system plus installed applications – publicly available. GNU/Linux, being the only free and open source operating system, is just a tool to do this. Success is, therefore, measured in terms of viability or practicability, as well as being able to help others, and not so much in terms of the overall number of users, or general adoption.

References

Arjona Reina, L. (2012) *Translations in Libre Software*. Madrid: Master Thesis, Universidad Rey Juan Carlos.

Baldwin, A. (2008) *Linguas OS – Linux for Translators: a Review of Linguas OS.* Available at: https://www.translatorscafe.com/cafe/EN/article85.htm [Accessed 20 June 2015].

Cronin, M. (2003) *Translation and Globalization*. London: Routledge.

Cronin, M. (2010) The Translation Crowd. *Revista Tradumàtica* 08/2010, 1-7. Available at: http://www.fti.uab.es/tradumatica/revista/num8/articles/04/04.pdf [Accessed 20 June 2015].

Diaz Fouces, O. (2010) *Faça você mesmo: GNU/Linux para tradutores (softwares livres e tradução)* Conferencia invitada promovida polo Departamento de Letras Modernas e o Centro Interdepartamental de Tradução e Terminologia da Faculdade de Filosofia, Letras e Ciências Humanas da Universidade de São Paulo.

Diaz Fouces, O. (2011) ¿Merece la pena introducir el software libre en la formación de traductores profesionales? In Universitat de Vic (ed.) *Anais das XI Jornadas de Traducción y Lenguas Aplicadas – Congreso Internacional "Didáctica de las lenguas y la traducción en la enseñanza presencial y a distancia" CDROM Language and Translation Teaching in FacetoFace and Distance Learning*. Facultat de Ciències Humanes, Traducció i Documentació de la Universitat de Vic.

Diaz Fouces, O. and García González, M. (eds.) (2008) *Traducir con software libre.* Granada: Comares.

DistroWatch.com (n.d.) Distrowatch.com – Put the fun back into computing. Use Linux, BSD. Available at: http://distrowatch.com [Accessed 20 June 2015].

Dotsub (n.d.) dotsub – The leading source for video translation and captions. Available at: https://dotsub.com/about [Accessed 20 June 2015].

EU Commission (2015) *Open Source Software Strategy 2014-2017*. Available at: http://ec.europa.eu/dgs/informatics/oss_tech/strategy/strategy_en.htm [Accessed 20 June 2015].

EU Commission (2015b) *Joinup – Share and reuse interoperability solutions for public administrations*. Available at: http://joinup.ec.europa.eu [Accessed 20 June 2015].

FLISoL (n.d.) *Festival Latinoamericano de Instalación de Software Libre.* Available at: http://www.flisol.info [Accessed 20 June 2015].

FOSS4Trans (n.d.) A Catalogue of Free/Open-Source Software for Translators. Available at: http://traduccionymundolibre.com [Accessed 20 June 2015].

FSF (n.d.) *Free Software Foundation*. Available at: http://www.fsf.org/ [Accessed 20 June 2015].

FSFE (n.d.) Free Software Foundation Europe: *Free Software and Education.* Available at: http://fsfe.org/freesoftware/education/education.en.html [Accessed 20 June 2015].

García González, M. (2013) Free and Open Source Software in Translator Education. The MINTRAD Project In *The International Journal for Translation & Interpreting Research*. 5-2, 125-148.

GETLT (n.d.) *Grupo de Estudos das Tecnoloxías Libres da Tradución.* Available at: http://webs.uvigo.es/getlt/ [Accessed 20 June 2015].

GTK (n.d.) The *Google Translator Toolkit*. Available at: https://translate.google.com/toolkit/ [Accessed 20 June 2015].

Hartley, Matt (2015) *Secret to Desktop Linux Adoption*, on-line page 2. Available at: http://www.datamation.com/open-source/secret-to-desktop-linux-adoption-2.html [Accessed 20 June 2015].

LinguasOS discussion group (n.d.) Available at: http://tech.groups.yahoo.com/group/linguasos/ [Accessed 20 June 2015].

Linguee (n.d.) English-German Dictionary. Available at: http://www.linguee.com [Accessed 20 June 2015].

Linux Counter Project (n.d.) Get counted as a Linux User. Available at: https://www.linuxcounter.net [Accessed 20 June 2015].

Linux Foundation (2015) The Linux Foundation. Available at: http://www.linuxfoundation.org/ [Accessed 20 June 2015].

Matecat (n.d.) Welcome to MateCat! Post-editing and outsourcing made easy. Available at: https://www.matecat.com/features/ [Accessed 20 June 2015].

McIlroy, M. D., Pinson, E. N. and Tague, B. A. (1978) Unix Time-Sharing System Forward In *The Bell System Technical Journal*, 1978, 57 (6, part 2). Bell Laboratories, 1902.

MyMemory (n.d.) at translated.net. Available at: http://mymemory.translated.net [Accessed 20 June 2015].

Net Market Share (n.d.) *Desktop Operating System Market Share*. Available at: http://www.netmarketshare.com/operating-system-market-share.aspx?qprid=8&qpcustomd=0&qpsp=2015&qpnp=1&qptimeframe=Y [Accessed June 2015]

Prior, M. (2003) Close Windows, Open Doors in *Translation Journal* 7(1). Available at: accurapid.com/journal/23linux.htm [Accessed 20 June 2015].

Prior, M. (2010) *Linux for Translators*. Available at: http://www.linuxfortranslators.org [Accessed 20 June 2015].

Sandrini, P. (2012) Translationstechnologie im Curriculum der Übersetzerausbildung In Zybatow, L. and Małgorzewicz, A. (eds.) *Sprachenvielfalt in der EU und Translation. Translationstheorie trifft Translationspraxis*. Neisse, 107-120. Available at: http://www2.uibk.ac.at/downloads/trans/publik/transtech-wroclaw.pdf [Accessed 20 June 2015].

Stallman, R. (2014) *Linux and the GNU System*. Available at: http://www.gnu.org/gnu/linux-and-gnu.en.html [Accessed 20 June 2015].

The Linux for Translators Forum (2002) Available at: https://groups.yahoo.com/neo/groups/linuxfortranslators/info [Accessed 20 June 2015].

The Rosetta Foundation (n.d.) *Trommons*. Available at: http://www.therosettafoundation.org/trommons/ [Accessed 20 June 2015].

tuxtrans discussion group (n.d.) Available at: http://groups.google.de/group/tuxtrans/ [Accessed 20 June 2015].

tuxtrans (n.d.) tuxtrans – GNU/Linux Desktop for Translators. Available at: http://www.tuxtrans.org [Accessed 20 June 2015].

Unicode website (n.d.) *The Unicode Consortium*. Available at: http://unicode.org/ [Accessed 20 June 2015].

W3Schools (n.d.) *OS Platform Statistics and Trends*. Available at: http://www.w3schools.com/browsers/browsers_os.asp [Accessed 20 June 2015].

Wikipedia (n.d.) *Usage share of operating systems*. Available at: https://en.wikipedia.org/wiki/Usage_share_of_operating_systems [Accessed 20 June 2015].

Zetzsche, Jost (2014) *The Translator's Toolbox. A Computer Primer for Translators*. International Writers' Group, LLC.

A Quality Model for the Evaluation
of Open Translation Technologies

Silvia Flórez, Amparo Alcina
Universidad de Antioquia, Colombia; Universitat Jaume I, Spain

1 Introduction

In recent years, like many other professions, translation has undergone a series of transformations as a result of the advances made in information and communication technologies. Since the beginning of the nineties the use of computer tools by translators has grown steadily, as has the number and variety of tools available, which range from general programs like text editors or processors to specific tools for translators such as translation memory systems (Alcina 2008). Faced with an ever-increasing array of tools to choose from, the translator is left wondering which of them would best fit his or her needs, often without the parameters required to be able to compare them and make an informed decision.

Now, although the area of technologies applied to translation has undoubtedly received a great deal of attention in the scientific and professional literature, it is also true that free and open source software has been largely neglected without being given the attention it deserves. The software we are dealing with here is characterised by guaranteeing the four fundamental freedoms for users described in the introduction to this volume.

Open software in general has advanced a great deal in recent years and new projects appear every day. Yet, according to the results of a study conducted by García (2008) to determine the situation of the translation technologies market, it would seem that most translators are unaware of and have little interest in the open software specifically designed for translation. Although García's study revealed that a good number of translators use open tools for tasks that are not related to translation, open translation memory systems are only just beginning to be considered as feasible options. In a profession in which the tools that have led the market for years cost hundreds of euros, the predominant popular conception seems to be that something that is free is not likely to be of good quality.

The question then arises as to how to make it easier for translators to identify the open programs that really do meet their needs. To obtain a possible answer to such a question we can resort to the criteria that have been used in the fields of software engineering and information systems, as well as in the specific area of translation technologies.

2 Evaluation of Software Quality

To begin with, we find that in software engineering quality is defined as "the extent to which an object (…) (e.g. a process, product or service) satisfies a series of specified attributes or requirements" (Schulmeyer 2006: 6). As regards the definition of the object, there are two different conceptions: one more restricted, known as small q, which comprises only the intrinsic product quality, and another more general one, known as big Q, which, in addition to taking the product into account, also covers the development process and user satisfaction (Kan 2002).

In practice, in recent decades two main approaches have been followed to understand and study software quality. One of them is diachronic and based on quality management, in which a flexible qualitative standpoint and a corrective methodology (normally used internally within the organisation that develops the software) are adopted. The other one is based on quality models, in which a descriptive methodology is followed with a more rigid perspective from which quality is understood as a quantifiable concept, either in terms of adherence to processes or based on the measurement or appraisal of a series of attributes (Groven et al. 2011).

The ISO 9126 standard ("ISO/IEC 9126. Software engineering. Product quality" 2001), which establishes a software quality model and guidelines for using that model, follows this second approach. This general-purpose quality model is made up of two parts: the first part specifies the characteristics that allow the internal and external quality of the software to be determined, while the second part deals with the concept of quality in use. The internal and external quality of the software as a product refers to the properties of the software itself and, according to the ISO standard, comprises six characteristics: functionality, reliability, usability, efficiency, maintainability and portability (see Figure 1). Quality in use, on the other hand, refers to the extent to which a given user can achieve his or her goals in a specific set of conditions of use. According to the ISO 9126 standard (2001), quality in use can in turn be broken down into four characteristics: effectiveness, productivity, freedom from risk and satisfaction (see Figure 2).

Another standard that also deals with software evaluation is ISO 14598 ("UNE-ISO/IEC 14598. Information Technology. Software Product Evaluation" 1998). This standard provides a general description of the software evaluation process and is therefore normally used in conjunction with the ISO 9126 standard.

In the field of translation technologies, software evaluation has a long history going back to the ALPAC report in 1966 on the status of machine

translation. Yet, given the abundance and diversity of tools and the variety of stakeholders and possible usage scenarios (industry, public administration, researchers, developers, agencies, freelance translators, students, etc.), there is a need for standard evaluation methods that are reliable, acceptable and reproducible (Quah 2006; Rico 2001; Höge 2002).
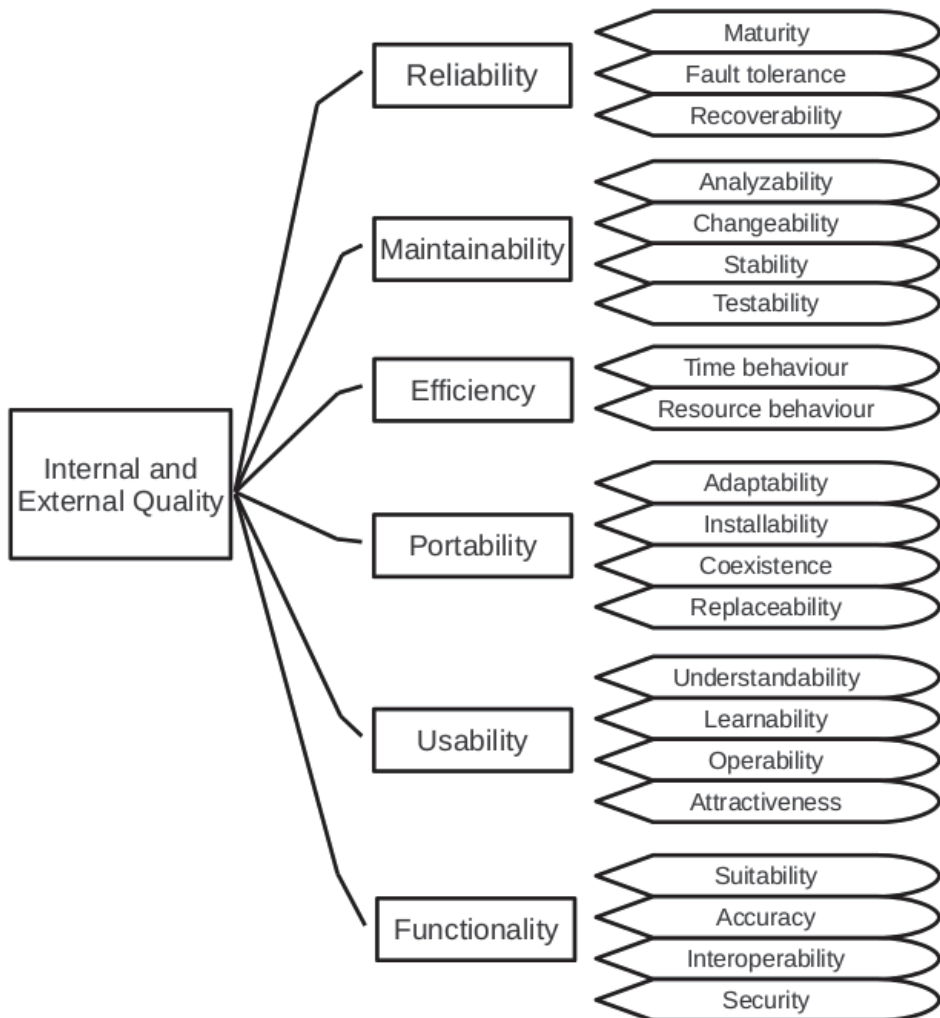
Figure 1: Internal and external software quality according to the ISO 9126 standard (2001).

As highlighted by Quah (2006), in the case of translation memory systems, evaluation is often part of the process of program development and is carried out from the point of view of researchers and developers rather than from that of the final user. Furthermore, in many cases the programs are evaluated by the same companies that develop them and, due to the fierce competition that exists in this field, the results are generally considered to be confidential.
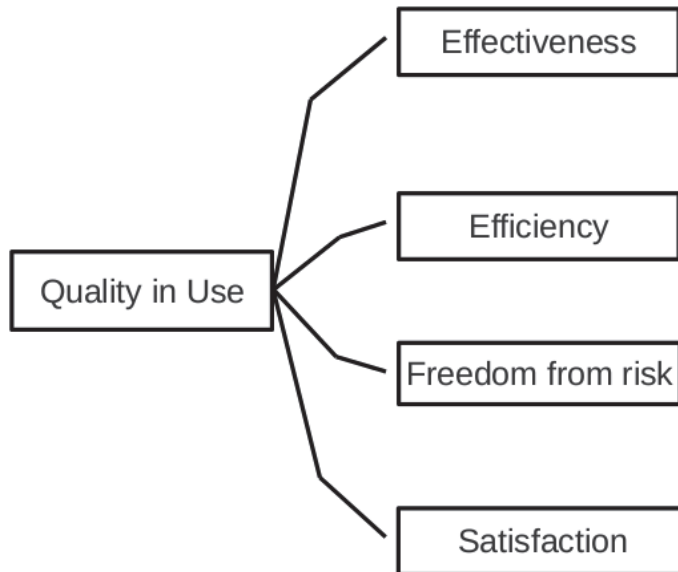
Figure 2: Quality of use according to the ISO 9126 standard (2001).

In an effort to find a solution to the problem of the lack of standardised evaluation criteria mentioned above, several attempts have been made to establish a general framework or series of reference guidelines for the evaluation of language technologies (Quah 2006), a category that encompasses translation technologies. The first of these initiatives was undertaken in 1993 by the Expert Advisory Group on Language Engineering Standards (EAGLES), funded by the European Union, and was based on the six quality characteristics proposed by the ISO 9126 standard.

Following the work carried out by EAGLES, in the year 2000 Europe and the United States began a joint project called International Standards for Language Engineering (ISLE). The project had three working groups, one of

which was devoted to the subject of evaluation (Evaluation Working Group, EWG) (Calzolari et al. 2003). The work of this group focused on the area of machine translation, as this is one of the most difficult technologies to evaluate, although the long-term idea was to be able to generalise the results obtained to the evaluation of other language technologies (Calzolari et al. 2003).

The work of this group resulted in the development of the Framework for the Evaluation of Machine Translation in ISLE (FEMTI), which is a structured collection of methods for evaluating machine translation systems (Calzolari et al. 2003; Quah 2006). Another work deriving from the EAGLES initiative was the Test-bed Study of Evaluation Methodologies: Authoring Aids (TEMAA), the main aims of which were to foster thought about the process of evaluating natural language processing tools and to work on the creation of a tool that was capable of carrying out that process automatically (Quah 2006; TEMAA n.d.). Within the framework of the project, case studies were carried out on the evaluation of spelling and grammar checkers, as well as information retrieval tools.

2.1 Evaluation of Translation Technologies

The theoretical model of the ISO 9126 and 14598 standards and the work by the EAGLES group have since given rise to several projects that include some kind of evaluation of translation technologies.

In her doctoral thesis, Höge (2002) presents her thoughts resulting from ten years of work in the field of translation technology evaluation from the user's point of view. Her work applies and complements the theoretical framework of the EAGLES group on the evaluation of different translation memory systems as part of the ESPRIT II project (1987-1992), financed by the European Commission. To apply her methodological proposal, the author evaluates two translation systems: Trados Translator's Workbench and IBM TM/2.

Rico (2001) also puts forward a final user-oriented model of evaluation that is based on the methodology proposed by EAGLES and the quality characteristics defined by the ISO 9126 standard. Her aim was to define a general model that could be re-used and applied in different translation contexts.

Maślanko (2004) conducted a comparative study of the terminological management modules integrated into a number of different translation memory systems (Multiterm iX by Trados, Déjà vu X by Atril and SDLX 2004 by SDL International). Her aim was to create an objective and detailed evaluation methodology that freelance translators and one-person translation businesses could use to select tools in Poland, her country of birth.

In her doctoral thesis, Filatova (2010) proposes adapting a scientific model of evaluation to the practical needs of translators. This project is broader in terms of the types of software evaluated, since it covers not only tools that, according to the author, are specific for translators (multilingual electronic dictionaries, word and character count, corpus analysis, translation memory suites) but also tools that she classifies as office automation software (file compressors, web browsers, e-mail clients, office automation suites, PDF readers and web authoring applications).

Finally, the work by Guillardeau (2009) is, according to the author himself and as far as we know, the first study to focus exclusively on the comparative evaluation of free translation memory systems. The author takes the quality criteria proposed by ISO and by the EAGLES group and the doctoral thesis by Lagoudaki (2008) on the functionality of translation memory systems as the basis for a qualitative comparison of two open tools (OmegaT and Anaphraseus) in terms of their functionality, efficiency and usability.

A number of works have addressed the evaluation of translation technologies but have been limited to very specific issues (such as Cerezo 2003; Gow 2003; and Lagoudaki 2007) or to providing simple comparisons of the functionality of the tools (such as, for example, the work by Zerfaß 2002; Bowker and Barlow 2004; Eisele et al. 2009; and Wiechmann and Fuhs 2006).

2.2 Evaluation of Free/Open-Source Software

As regards the quality of free software, in recent years the fields of software engineering and information systems have adapted evaluation methodologies that take into account the specific features of this type of software and its development paradigm. In addition to evaluating the software as a product, they also cover aspects related with the communities that support the projects (Samoladas et al. 2008).

The first specific quality models, which appeared between 2003 and 2005, are known as first-generation models and are based on the traditional quality models of proprietary software, but have been adapted and complemented so as to make them applicable to free software (Groven et al. 2011). Some of the more notable first-generation models include the Open Source Maturity Model (OSMM) developed by Capgemini in the year 2003, the OSMM developed by Navica in 2004, one developed by the project Qualification and Selection of Open Source Software (QSOS) (Atos Origin 2006), originally started by Atos Origin in 2004, and the project Business Readiness Rating (BRR) (BRR 2005; Wasserman et al. 2006), which was begun by the Carnegie Mellon West Centre for Open Source Investigation and Intel, among others, in the year 2005 (Groven et al. 2011).

The quality models for free software that have appeared since 2006 are known as second-generation models and are characterised by being based on both the traditional models of proprietary software and on the first-generation models. Moreover, they are focused on the automation of the evaluation process and on providing more advanced metrics and tools for evaluation that are made available as web applications or plug-ins for development environments (Groven et al. 2011). Some of the better-known second-generation quality models include those developed by the projects Quality in Open Source Software (QualOSS) (Deprez 2009), Quality Platform for Open Source Software (QualiPSo) (Wittmann and Nambakam 2010), and Software Quality Observatory for Open Source Software (SQO-OSS) (Samoladas et al. 2008), all of which were funded by the European Community (Deprez and Alexandre 2008).

3 Towards a Method of Evaluation for Open Translation Technologies

In this context, an objective detailed evaluation of the open tools for translators currently available may be a good way to disseminate the concept of free software in our profession and foster its use. The evaluation methods traditionally used for language technologies are focused on sequential or iterative and incremental development cycles and design processes rather than on non-continuous cycles such as those of free software (Gasser, and Scacchi, Ripoche and Penne 2003). Hence, there is a need for an integral evaluation methodology which takes into account not only the software as a final product but also considers aspects related to the development project, such as intellectual property management, forward planning, the dynamics of the user and developer communities, and the technologies supporting them.

In this work we therefore propose a method for evaluating open translation technologies. The method outlined here comprises a quality model and guidelines for its use (the activities, tasks and participants in the evaluation process, and the expected use of the results).

Taking an interdisciplinary perspective that includes technological, sociological and business aspects as our starting point, a qualitative approach was adopted for the evaluation. The reason underlying this decision was that the main interest was to describe the characteristics of the ecosystem of open translation technologies and to explore the feasibility of the programs currently available, rather than to reach generalisations about this type of software. The aim of the proposed method is to help translators when it comes to choosing open tools to integrate within their work environment. The users of

the results of the evaluation are expected to be freelance translators, trans-lation teams, small companies, researchers, and translation students and teachers interested in open translation technologies.

3.1 Activities and Steps of the Evaluation

The method of evaluation proposed here comprises three main activities that are in turn divided into a series of steps, as detailed in the following:

- Preparing the evaluation: this consists in defining the type of tests and the quality model (the categories and criteria to be taken into account and the metrics and procedures for consolidating the results) and in designing and implementing the instruments.
- Evaluation: this consists in determining the sample of projects to be evaluated and collecting data by applying the questionnaire, which automatically generates the records with the results.
- Selection: this consists in specifying the user's requirements (existing environment, work formats and functional modules depending on the tasks to be undertaken), comparing programs that meet those require-ments and choosing the most suitable.

In this case the first two activities of the process are carried out by the researcher herself, whereas the final or selection phase is to be done directly by the final user. In the following, we will concentrate on detailing the first of these activities, that is to say, on preparing the evaluation. For illustrative purposes, we will present the results of the evaluation of the open translation memory system OmegaT, which was conducted in May 2012.

It should be noted that this work was part of the research carried out by Flórez (2013) for her doctoral thesis, which included the compilation of a cata-logue of free/open-source software for translators (see Flórez and Alcina 2011a and 2011b), and the evaluation of eleven development projects working on desktop translation memory systems available under free licen-ces. Both the catalogue of tools and the instruments and full results of the evaluation are available in an online wiki created as part of that project (see Flórez 2012a).

3.2 Quality Model

To define the software quality model, the first step was to establish the type of test to be used and the context of evaluation. Bearing in mind that the rationale behind the evaluation of the software in this case was to test the general characteristics of the programs for their possible implementation in the translator's work environment, we decided to use the type of tests called feature inspection, the role of which is only to indicate the presence or

absence of certain features and not to identify bugs in the programs (EAGLES 1996; Höge 2002). This kind of tests was chosen because of its descriptive nature and due to the fact that it is simple, fast and easy to apply, since the data needed can be largely obtained from the documentation of the programs and the websites of the projects.

3.2.1 Categories and Criteria

In the hierarchy for defining the evaluation criteria we started out by drawing a distinction between project and product. The quality model is made up of two parts: the first allows the development projects to be characterised so as to gain a better understanding of the practices and processes involved, as well as the resources and services available to the community of users. The second refers to the quality of the software as a product and makes it possible to determine the features and technical characteristics of the programs.

Project Quality

With the aim of characterising the free translation technology development projects, based on what was found in the literature and following the re-commendation to work from the most general to the most specific, four characteristics were included: strategy, community, maturity and reputation of the project. Project quality is broken down into characteristics and sub-characteristics in Figure 3.

Product Quality

Taking into account the rationale behind the evaluation and the functional orientation of the programs, three of the six characteristics proposed in the ISO 9126 standard were used as criteria for evaluating the software, namely: functionality, usability and portability. Given the scope of this project, the other three characteristics set out in the ISO 9126 standard (reliability, maintainabi-lity and efficiency) were not included in the model. Figure 4 shows the characteristics and sub-characteristics of product quality that were included in the quality model.
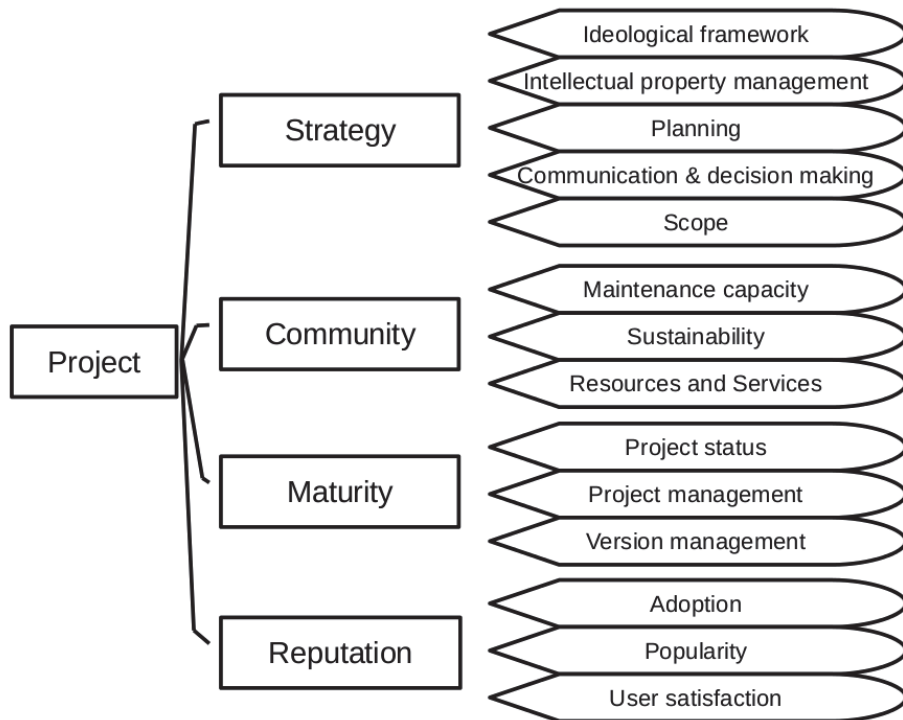
Figure 3: Characteristics and sub-characteristics of the quality of the project.
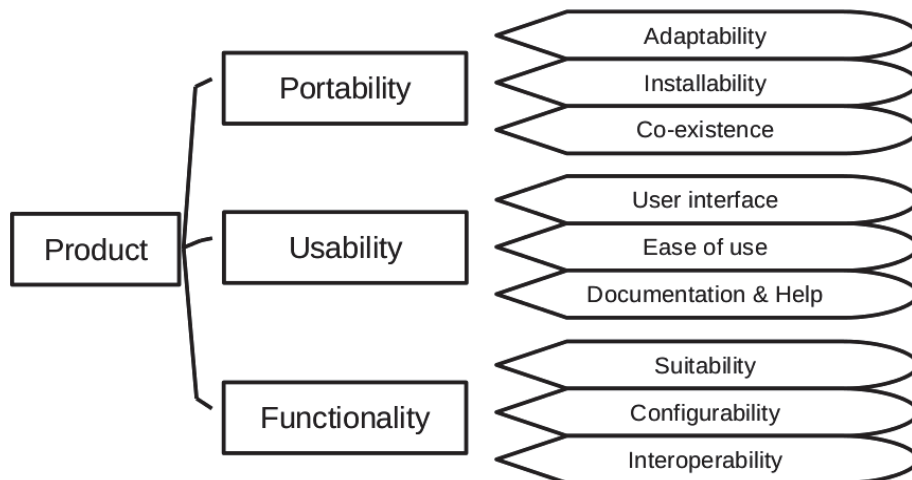
Figure 4: Characteristics and sub-characteristics of the quality of the product.

At this point it is important to note that the attributes corresponding to portability and usability are equally significant for any type of tool. In other words, they are non-functional criteria that can be applied both to a web browser and to an office automation application or to a translation tool. The attributes of the functionality characteristic, in contrast, vary according to the type of tool to be evaluated and the tasks that can be done with it (alignment, translation, proofreading, invoicing, etc.). It must be made clear that the quality model prepared for this study is limited to analysing the functionality of desktop translation memory systems.

3.2.2 Attributes and Metrics

The next step consisted in breaking down each of the quality characteristics and sub-characteristics into one or more attributes. In the case of the project quality characteristics, a qualitative assessment was chosen. This means that for these attributes no quantitative scores were defined; in contrast, the factual information is presented directly on the result sheets so that the users can broaden their knowledge on each project. For the non-functional characteristics of product quality (portability and usability), on the other hand, we defined the corresponding attributes and metrics, that is, the way to obtain the quantitative scores and the scales to be used in each case. Finally, for functionality, a checklist was drawn up where the characteristics that were present could be indicated, but neither scoring was used nor were any appraisals made about the features implemented.

Project Quality

The tables below show the attributes defined to evaluate the strategy (Table 1), community (Table 2), maturity (Table 3) and reputation of the projects (Table 4) and the possible answers established for each attribute. As can be appreciated in the tables, some attributes are binary (presence/absence), while others are classificatory and still others are numerical.

Project strategy		
Sub-characteristic	**Attribute**	**Options**
Ideological framework of the project	Origin of the project	Independent project Publicly funded project Privately funded project Mixed funding project
	Type of ethics that govern the project	Hacker ethics Hybrid ethics Business ethics

Project strategy		
Sub-characteristic	**Attribute**	**Options**
Intellectual property management	General licensing strategy	One free licence Several free licences Dual licensing (free/proprietary) Open core
	Permissiveness of the licence	Without copyleft With weak copyleft With strong copyleft
	Guidelines or transfer of rights agreements for collaborators	Presence Absence
	Ownership of copyright	The owner is a single developer Ownership assigned to a legal body Distributed ownership
Forward planning	Specification of requirements	Presence Absence
	Roadmap	Presence Absence
	Description of new anticipated features	Presence Absence
	Versions planning	Presence Absence
Communication and decision-making structures	Type of process for decision-making	Decentralised Balanced Centralised
	System of governance	Benevolent dictatorship Meritocracy Democracy Anarchy
	Mechanism of representation used by the project to communicate and be identified	Original developer Recognised leaders Foundation Steering committee Sponsoring institution or company

Project strategy		
Sub-characteristic	**Attribute**	**Options**
Scope	Integration of code from other free projects	Yes No
	Project derived from another free project	Yes No
	Development of other tools	Yes No

Table 1: Attributes to determine the project strategy.

Community		
Sub-characteristic	**Attribute**	**Options**
Maintenance capacity	Type of development community	Independent developer Group of developers Formally organised developers Legal body Commercial body
	Forks or derived tools	Presence Absence
	Institutions linked to the project	Presence Absence
	Number of active developers	Numerical value
	Number of subscribers in the lists of users	Numerical value
Sustainability	Number of users who participated in discussions over the last month	Numerical value
	Average number of messages per month in the users' forum in 2011	Numerical value
	Average response time in the forums (last 5 questions)	Numerical value
Resources and services available	Web portal highlighting significant information about the project	Presence Absence
	Communication spaces actively used in the last year (mailing lists, wikis, blogs, IRC chats, social networks)	Presence Absence

Community		
Sub-characteristic	**Attribute**	**Options**
	Personalised technical support	Presence Absence
	Added value subscriptions	Presence Absence
	Training (tutorials, video channel, webinars, etc.)	Presence Absence
	Personalised development	Presence Absence
	Consultancy	Presence Absence
	Software as a service	Presence Absence

Table 2: Attributes for characterising the project community.

Maturity of the project		
Sub-characteristic	**Attribute**	**Options**
Project status	Date the project started	Numerical value
	Current development status	Beta Stable Mature Inactive
Project management	Management of the project in one of the main public forges	Presence Absence
	Source code repository with revision tracking system	Presence Absence
	System for managing potential bug reports	Presence Absence
	System for managing new feature requests	Presence Absence
	Existence of documented processes to contribute to the project	Presence Absence
	Platform for managing the localisation of the program and the documentation	Presence Absence

Maturity of the project		
Sub-characteristic	**Attribute**	**Options**
	Documented process of eliciting and managing requirements	Presence Absence
Version management	Defined release cycle	Presence Absence
	Versions released in 2011	Numerical value
	Minor updates released in 2011	Numerical value
	Date of last version released	Numerical value

Table 3: Attributes for determining the maturity of the project.

Reputation of the project		
Sub-characteristic	**Attribute**	**Options**
Adoption	Books, publications, reviews or entries in blogs about the project	Presence Absence
	Reference implementation/success cases documented on the project website	Presence Absence
	Average number of downloads during the week following the release of the last three versions	Numerical value
Popularity	Number of downloads in the last month	Numerical value
	Discussions in translators' forums (ProZ, LinkedIn, etc.)	Presence Absence
	Packages included in GNU/Linux repositories	Presence Absence
	Project included in software catalogues or directories	Presence Absence
	Profile of the project on Ohloh.net	Presence Absence
User satisfaction	Reviews and scores in the forge used	Presence Absence
	Comments on the project on social networks	Presence Absence

Table 4: Attributes for determining the reputation of the project.

Product Quality

For the non-functional characteristics (portability and usability) of the software as a product, each sub-characteristic was broken down into a series of attributes, and then a series of possible answers and their associated scores were formulated for each of them. For these two characteristics we decided to use a homogeneous scale ranging from 1 to 3, where 1 means unacceptable, 2 is acceptable and 3 is satisfactory. While drafting the possible answers, efforts were made to consider the situations that are found in real use cases and special attention was paid to avoiding ambiguity, in an attempt to reduce the possibility of different interpretations being made by different evaluators in different contexts. Due to space restrictions, not all the attributes of these two characteristics are detailed here. For illustrative purposes, Table 5 below presents the possible answers for two attributes of portability and Table 6 shows two usability attributes.

Portability		Scoring		
Sub-characteristic	Attribute	1	2	3
Adaptability	Modularity	The design of the tool does not allow for the development of independent components.	The design of the tool allows for the development of independent components that can be integrated within the system, but no documentation is available.	The design of the tool allows for the development of independent components by means of a plug-in architecture or a well-documented public API.
	Scalability	The system is not designed with large-scale implementations in mind and does not include a multi-user mode.	The system can be implemented on a large scale, but it is not designed for multi-user environments or vice versa.	The system can be implemented on a large scale and in multi-user environments.

Table 5: Details of two attributes for evaluating the portability of the product.

Usability		Scoring		
Sub-characteristic	Attribute	1	2	3
User interface	Layout of the user interface	The interface is complex with too much information that is not clearly organised; the manual has to be used.	It takes some time to understand the interface, the information is more or less organised; the manual has to be used from time to time.	The interface is simple and intuitive, the information is well-organised; the manual is practically not needed.
	Availability of the required language	The program and its documentation and help are only available in a language other than the one required.	Localisation is partial (interface in the required language but documentation is not translated or vice versa).	The programme is totally localised into the required language, including both the user interface and the help, as well as other documentation that is included.

Table 6: Details of two attributes for evaluating the usability of the product.

In order to evaluate functionality, the features included, the possible configurations, the capacity to process different input formats and the interoperability were considered. A checklist was established with the main characteristics that one can expect to find in translation memory systems based on the functional descriptions of the principal commercial proprietary systems and on previous knowledge about this kind of tools. Following this same line, the list of features and supported formats can easily be expanded to cover other types of programs.

For each of the functionality attributes the presence or absence of the characteristic in question is indicated, but no scores are calculated and the adequacy of feature implementation is not appraised. In contrast, the full list of characteristics present is included on the result sheet. Table 7 offers details of the attributes that were used to evaluate the functionality of the programs belonging to the type translation memory systems.

Functionality		Scoring
Sub-characteristic	**Attribute**	**Presence or absence**
Suitability for purpose	Match between the features included and the expected features according to the type of program	Project options: Analysis of originals (wordcount, matches, repetitions) Batch processing Pre-translation of documents Pre-translation prioritising the sources used Pseudotranslation Creation of projects with multiple source documents Possibility of using the memories in both directions Multiple memories per project Multiple glossaries per project Multiple translations for the same original segment Multilingual memories (more than two languages) Simultaneous use of glossaries/memories shared over the web Fuzzy matches Context-based matches Glossary matches Automatic insertion of exact matches Automatic insertion of fuzzy matches Automatic propagation of repeated segments
		Editor options: Visualisation of metadata of the matches (date, user ID, project, etc.) Segment validation by means of different statuses Option of browsing around the editor by means of filters Possibility of adding comments to the segments Project statistics (number of segments translated/not translated) Global search and replace Search for concordances in original files Search for concordances in reference files On-the-fly auto-complete

Functionality		Scoring
Sub-characteristic	**Attribute**	**Presence or absence**
		On-the-fly spellchecker On-demand spellchecker On-the-fly grammar/style checking On-demand grammar/style checker Preview of format Review mode (track changes, comments, export to table) On-the-fly quality checks On-demand quality checks
		Integration with external applications: Integration with local or web-based machine translation engines Search in external resources (local or via web services) Integration with voice recognition software (commands and/or dictation)
	File filters implemented	Text and office automation formats: TXT, CSV, TAB, DOC, DOT, RTF, XLS, XLT, PPT, PPS, DOCX, DOTX, XLSX, XLTX, XLSM, PPTX, PPSX, POTX, ODT, ODS, ODP, SRT
		DTP formats: MIF (FrameMaker), XML (FrameMaker), INX (InDesign), IDML (InDesign), tagged TXT (Pagemaker, Ventura), QSC (QuarkXPress), XTG (QuarkXPress), TTG (QuarkXPress), TAG (QuarkXPress), IASCII (Interleaf/QuickSilver), PDF (Acrobat Reader)
		Multimedia formats: PSD (Photoshop), SVG (Photoshop, Illustrator, CorelDraw, generic), DXF (AutoCAD), TXT (AutoCAD)
		Web localisation formats: HTML, XML, ASP, PHP, JSP, INC, NET, RESX, PPSM, XAML, SGM
		Software localisation formats: RC, DLG, EXE, DLL, MO, PO(T), Java Resource Bundles, XML (Android resource), XIB (iOS App resource), TS (Qt Linguist), QPH (Qt Phrase Book), DTD (Mozilla)

Functionality		Scoring
Sub-characteristic	**Attribute**	**Presence or absence**
Configurability	Possibility of configuring the system according to different needs	Configurable filters Configurable segmentation rules Possibility of changing segmentation during translation Configurable minimum percentage of matches Customisable spellchecker dictionaries Customisable language corrector rules Searches and replacements based on regular expressions Configurable placeables and localisables (dates, variables, etc.) Configurable quality checks (tags, punctuation, spaces, numbers, terms, etc.) Control of access to the system by means of users and permissions Configurable keyboard shortcuts
Interoperability	Support for data exchange standards	Unicode encoding SRX segmentation rules TMX memories TBX databases Glossaries as delimited text (CSV, TAB or TXT) Pre-translated XLIFF files
	Support for open formats generated by other translation tools	TTX (SDL Trados) TXT (WordFast) TXML (WordFast Pro) NXT (STAR Transit)

Table 7: Attributes for evaluating the functionality of the product.

3.2.3 Procedures for Consolidating the Results

Procedures were then defined for summarising the attribute data in global scores per sub-characteristic. Since it was a general exploratory evaluation, all the attributes and characteristics were considered to be of equal importance and we therefore decided not to weight the results because we did not set out from a specific evaluation context that justifies the assignation of particular values. Moreover, the use of different scales (binary, classificatory and ordinal) makes weighted averages unsuitable for the consolidation of results.

As regards the quality of the project, for the characteristics project strategy and reputation we decided not to summarise the results by means of indicators as these aspects were not considered to have a decisive effect on the selection of the tools. In contrast, the information about the project strategy is presented on the result sheets as a descriptive paragraph about the projects, whereas the data found about their reputation is included as reference links for those interested in such information.

The results of the other two characteristics of project quality (community and maturity) were summarised by defining the acceptance criteria shown in Table 8. If the project met the established criteria, a star was given for the corresponding sub-characteristic; the project can thus obtain a maximum of three stars per characteristic. The number of stars obtained is interpreted as follows: 3 stars = satisfactory, 2 stars = acceptable, 1 star = poor, 0 stars = unacceptable. Furthermore, it was decided that for the projects with no stars for the characteristics of community and maturity the software would not be evaluated as a product.

Characteristic	Sub-characteristic	Acceptance criteria
Community	Maintenance capacity	At least one active developer and a users' forum with subscribers.
	Sustainability	Existence of active discussions in the last month and an average of no fewer than four messages per month over the last year.
	Resources and services available	Web portal with relevant information about the project; at least two communication spaces where users can obtain answers to their doubts.
Maturity	Project status	The project must be at least two years old and its current development status must be stable or mature.
	Project management	The code must be managed in a public forge with a change tracking system and bug report management.
	Version management	The project must have released at least one version or update in 2011 and the latest version available must be from 2011 or 2012.

Table 8: Indicators of the quality of the community and maturity of the project.

As to product quality, for the non-functional characteristics (portability and usability) stars are also assigned per sub-characteristic, but in this case the procedure used to obtain the global scores consists in simply adding up the individual scores of the attributes of each sub-characteristic and classifying the results in accordance with Table 9.

Lastly, for functionality, the information is not consolidated but instead, as explained in the previous section, the list of features available and the file formats supported are presented on the result sheet.

Characteristic	Sub-characteristic	Acceptance criteria
Portability	Adaptability	Minimum score equal to or higher than four.
	Ease of installation	Minimum score equal to or higher than six.
	Coexistence	Minimum score equal to or higher than four.
Usability	User interface	Minimum score equal to or higher than eight.
	Documentation	Minimum score equal to or higher than six.
	Ease of use	Minimum score equal to or higher than eight.

Table 9: Quality indicators for portability and usability.

3.2.4 Evaluation Instrument

The evaluation instrument was implemented as a complement to the catalog of open-source software for translators available in an on-line wiki created specifically for this purpose (see Flórez 2012a). Thus, we have a repository that makes both the instruments and the evaluation results publicly available. The instrument enabling the evaluator to collect data consists in a series of web forms (one for each quality characteristic, see Figure 5) that are filled in by hand. The data obtained are presented as complementary information on the data sheets in the catalogue.

Figure 5: Evaluation instrument – Project strategy.

4 Results

Below, we present the results of the characterisation of the OmegaT develop-ment project and the evaluation of the tool broken down by characteristics.

4.1 Characterisation of the Project

In the following subsections we present the results for each of the sub-charac-teristics of the quality of the OmegaT project, namely strategy, community, maturity and reputation.

4.1.1 Strategy

The OmegaT project began as an initiative by independent developer Keith Godfrey and now has a group of recognised leaders. The work is carried out on a voluntary basis. The software and its features are available under a GNU GPL (strong copyleft) license and ownership is distributed among its developers. According to the philosophy of the project stated on its website it

is a "delegated anarchy", where anyone is free to contribute to the project and there is a central team of developers who decide what contributions are to be included in the code that is distributed to the community. The project integrates code developed by other free projects (Hunspell, LanguageTool, Lucene Tokenizers and Okapi Framework).

4.1.2 Community

The project has a website (http://www.omegat.org) where relevant information is posted. Development is carried out by a group of developers in a collaborative and informal manner. In March 2012 there were four active developers and the user group had 1720 subscribers, of whom 39 had participated actively in the previous month. Moreover, the project has a general manager, a development manager, a documentation manager and a localisation manager.

In 2011 there were an average of 304 messages per month in the user group and the average response time for the last five questions was 0.3 hours; it is not necessary to be a member of the group to consult the message archive. The project also has a mailing list for developers and another for localisation management. In addition, it has an IRC channel. With regard to the services it offers, it is possible to sponsor the development of new features by getting in touch with the developers directly in order to agree upon the value of the monetary contribution to be paid.

There are several projects derived from OmegaT, some of the more important being: OmegaT+, a fork started by one of the developers following a series of disagreements (at the time of writing there are still disputes between the two projects over the name OmegaT as the trademark registered by the original project); Boltran, a web-based version of OmegaT; and Autshumato ITE, a translation memory system that integrates OmegaT, OpenOffice.org and the machine translation engine Moses (in this case there is some degree of collaboration between the projects).

Appraisal

In this case the fact that there is a website which is both well organised and offers detailed information about the project is judged positively, as are the number of active collaborators and the existence of derived projects. Furthermore, another positive point is the existence of several communication spaces for members of the project, together with the level of activity and the response time in the users' forum. As regards the professional services on offer, although the possibility of sponsoring the development of new features is valued positively, bearing in mind the characteristics of the project there could be a greater range of professional services on offer.

4.1.3 Maturity

According to the copyright statement, the project began in the year 2000 and was registered on SourceForge on November 28th 2002. The current development status is stable and two main parallel versions are maintained: the so-called standard version, with all the features duly documented, and another called latest, previously known as beta, which the developers claim is equally stable but differs from the first one in that the latest features are not yet documented and the localisation may not be totally up-to-date. In 2011, 2 main versions and 13 updates were released and at the time of the evaluation (May 2012) the most recent standard version (2.5.4) was from May 9th 2012.

The project uses a repository with revision tracking (SVN) for code management and the tools provided by SourceForge for bug management and new feature requests. There is a documented process for contributing with the localisation of the interface and the documentation of the program.

Appraisal

The age of the project and its current development status are valued positively, as is the use of a public forge and specific tools for code management, bug reports and new feature requests. Furthermore, although there is no predefined release cycle, the regular release of updates and the availability of a recent version are given a positive appraisal.

4.1.4 Reputation

In March 2012 the software was downloaded 5033 times and the average number of downloads carried out during the week following the release of the latest three versions was 1344, a figure which can be used to get an idea of the number of regular users of the tool. A number of publications about OmegaT were found and specific discussions were observed in translators' forums, for example, the support group in ProZ and a group in LinkedIn called OmegaT Translation Professionals. OmegaT is also included in the repositories of several GNU/Linux distros and is listed in several software directories. According to the scores on SourceForge at the time the evaluation was carried out, 88% of users recommend the tool (170 recommendations versus 23 negative ratings). Recent comments were also found on Twitter and the project has an updated profile on Open Hub (previously known as Ohloh), a platform for free software developers and projects where source code repositories of the programs are analysed and summaries of statistics are offered (including lines of code, programming languages and licenses used, level of activity of the projects and their estimated monetary value).

Appraisal

The existence of publications about the project and the high number of downloads are valued positively. Another positive point was the existence of discussions about the tool in translators' forums and its being included in software directories and GNU/Linux distros. Likewise, the existence of recommendations in the forge and comments on Twitter was valued positively, as was the updated profile on Open Hub.

4.2 Evaluation of the Software as a Product

As mentioned earlier, the OmegaT project maintains two parallel versions of the tool: the standard and the latest. The standard version was used for the evaluation of the product as it is the one recommended for users who are beginning to use the tool. At the time of the evaluation (May 2012), the standard version that was available was 2.5.4.

Here we include the results for the functionality of OmegaT (see Table 10). Portability and usability of the tool were also evaluated, but due to space restrictions they are not included here; the detailed results of these two characteristics can be consulted in Flórez (2012b).

Functionality		
Sub-characteristic	**Attribute**	**Characteristics present**
Suitability for purpose	Match between the features included and the expected features according to the type of program	Project options: Analysis of originals (wordcount, matches, repetitions) Batch processing Pre-translation of documents Pre-translation prioritising the sources used Pseudotranslation Creation of projects with multiple source documents Fuzzy matches Context-based matches Automatic insertion of exact matches Automatic insertion of fuzzy matches Automatic propagation of repeated segments Glossary matches Multiple glossaries per project Possibility of using the memories in both directions Multiple memories per project

Functionality		
Sub-characteristic	**Attribute**	**Characteristics present**
		Multiple translations for the same original segment Multilingual memories (more than two languages)
		Editor options: Visualisation of metadata of the matches (date, user ID, project, etc.) Option of browsing around the editor by means of filters Possibility of adding comments to the segments Project statistics (number of segments translated/not translated) Search for concordances in original files Search for concordances in reference files On-the-fly spellchecker On-the-fly grammar/style checking On-demand quality checks
		Integration with external applications: Integration with local or web-based machine translation engines
	File filters implemented	Text and office automation formats: TXT, CSV, TAB, DOC, DOT, RTF, XLS, XLT, PPT, PPS, DOCX, DOTX, XLSX, XLTX, XLSM, PPTX, PPSX, POTX, ODT, ODS, ODP, SRT DTP formats: XML (Infix), IDML (InDesign), XTG (QuarkXPress), TAG (QuarkXPress) Multimedia formats: SVG, XML (Flash export), CAMPROJ (Camstasia Studio) Web localisation formats: HTML, XML, RESX, JSON Software localisation formats: RC, POT, PO, Java Resource Bundles, XML (Android resource), TS (Qt Linguist), DTD (Mozilla), HHC (HTML Help Compiler)
Configurability	Possibility of configuring the system according	Configurable filters Configurable segmentation rules Configurable minimum percentage of matches

Functionality		
Sub-characteristic	**Attribute**	**Characteristics present**
	to different needs	Customisable spellchecker dictionaries Customisable language corrector rules Searches based on regular expressions Configurable quality checks Configurable keyboard shortcuts
Interoperability	Support for data exchange standards	Unicode encoding TMX memories TBX databases Glossaries as delimited text (CSV, TAB or TXT) Pre-translated XLIFF files
	Support for open formats generated by other translation tools	TXML (WordFast Pro)

Table 10: Functionality of OmegaT (2012).

Table 10 shows the characteristics offered by OmegaT, version 2.5.4. As can be observed, the list of features included and formats supported is quite extensive and covers the most common requirements for exchanging data in our industry: Unicode, TMX, TBX and XLIFF. It should be noted that some features that were not available at the time of the evaluation (e.g. the search and replace option within the project) have since been implemented in later versions of the tool. Furthermore, the possibility of adding functionality by means of scripts (which were previously available as a plug-in and from version 3.0.3 onwards as a built-in feature) means that OmegaT can be adapted to the specific requirements of the translator's workflow.

4.3 General appraisal

The general appraisal is established by combining the appraisals of the characteristics that have been evaluated. The fact sheet of the general appraisal of OmegaT is available in the wiki, as can be seen in the partial screenshot presented in Figure 6. Owing to space restrictions, the list of features and supported formats has been excluded as this information was already shown in Table 10. As can be seen in the figure, according to the data obtained, both the community and maturity of the OmegaT project and the

portability and usability of the tool are considered satisfactory (three stars). The fact sheet also provides information about the strategy of the project, as a descriptive paragraph, and about the reputation of the project, including links to the main resources related to it.

Figure 6: Partial screenshot of the fact sheet of the general appraisal of OmegaT.

5 Discussion

The evaluation instrument was tested with a sample of eleven open-source projects working on desktop translation memory systems; here we present the results for the OmegaT project. In our opinion, the results obtained allow possible users to make inferences about the project evaluated, to compare

them and to select the tool that is best suited to their needs. Additionally, in general terms, the results obtained are considered to reflect the characteristics of the projects evaluated and can help translators to familiarise themselves with the characteristic aspects of the free software that they should take into account when it comes to choosing a tool for their work environment.

Bearing in mind the exploratory approach followed in this work, in general terms the test evaluation has been positive. As a favourable aspect, the instrument can easily be updated to include new features if and when necessary.

During the evaluation process, however, we also detected several possible problems and aspects that could be improved in order to achieve a more rigorous and detailed evaluation. On evaluating the project strategy, for example, for the attributes type of process for decision-making (decentralised, balanced or centralised) and system of governance (benevolent dictatorship, meritocracy or anarchy), the explicit information needed was found on the websites of the projects in only one case. It is therefore clear that these two attributes are more complex than expected and so it would be recommendable to use other techniques to evaluate them, such as a detailed analysis of the archives of the mailing lists or interviews with the developers.

One aspect of the strategy of the projects that was not taken into account and that could help to improve our understanding of the scope of the project is the target users. Some projects, especially in the field of natural language processing, are aimed at users with an advanced knowledge of computers and developers who are used to working on command lines, that is, without graphic interfaces. In other cases the tools are web-based and are not offered as a service, which implies that their installation and maintenance lie beyond the possibilities of users whose technical know-how is limited to the desktop environment. It would therefore be useful to add the attribute target users as part of the sub-characteristic scope of the project, so that these data can be used to filter the tools, according to the technical know-how needed to use them.

With regard to the characterisation of the communities, the breakdown of the sub-characteristic sustainability could be improved. In the method proposed here, three attributes were employed: the number of participants in the user lists in the last month, the average number of messages per month in 2011 and the average response time for the last 5 questions asked in the forums. Nevertheless, the data needed to evaluate this last attribute were found for only two projects.

On evaluating the maturity of the projects, two attributes were considered as part of the sub-characteristic project status: the date the project began and

the current development status. In both cases the data were obtained from the development forges, but in some cases discrepancies were found between the self-classification by the projects themselves and the classification of the forge. Moreover, it is also necessary to take into account that free projects may change development forge, and therefore the date that appears may be at odds with the date the project was initially registered. This information should therefore be confirmed using other sources, such as the information provided by the websites and blogs of the project or the change log that is sometimes included in the downloads.

The evaluation of the reputation of the project is another aspect that could be dealt with in greater depth. This could be achieved using qualitative techniques, such as the analysis of contents posted in translators' forums and social networks, or surveys carried out on users in order to determine their degree of satisfaction with the tools.

As regards the portability of the tools, in order to calculate the time needed to install them, which is covered by the sub-characteristic ease of installation, the instrument could be improved by specifying that this refers to the basic installation of the tool, without including dependencies, plug-ins or add-ons. Furthermore, in order to evaluate the possibility of integrating the tools into the existing workflow, an attribute that corresponds to the sub-characteristic coexistence, the type of test used (feature inspection) may not be sufficient and it would be recommendable to go deeper into the evaluation of this aspect by means of scenario testing within the expected environment of use.

According to our findings, the evaluation of the usability of the tools is perhaps the characteristic that entails the greatest risk of subjectivity. Aspects of the user interface, such as the user-friendliness of its layout or how easy it is to understand the icons and features, largely depend on the evaluator's point of view and perhaps also on his or her degree of familiarity with the type of tools being evaluated. For example, for a translator who is used to working with segments in columns, a horizontal layout may seem less user-friendly and vice versa.

For the sub-characteristic ease of use, on the other hand, although the attributes appraised are of a more objective nature (possibility of browsing and operating with just the keyboard, existence of contextual help and the existence of progress indicators and error messages), more rigorous results could be achieved by using systematic menu-oriented tests, designed to examine all the features offered by a program sequentially.

6 Conclusions

In this chapter we present a quality model for the evaluation of open source translation technologies. The model proposed here was implemented in a wiki as a complement to a catalogue of free software and it was tested with eleven free projects working on desktop translation memory systems. Both the evaluation instruments and the results of the eleven projects evaluated are publicly available in a wiki. In our opinion the quality model can be useful, and the results can be of use to translators interested in free software, since the fact sheets that are generated allow them to view the basic information about the project and the tools. We believe that having this kind of information available in a public repository can make it easier for freelance translators to reach a decision when it comes to selecting free tools for their work environment.

References

Alcina, A. (2008) Translation Technologies: Scope, tools and Resources. Target: International Journal on Translation Studies, 20(1), 79-102.

Atos Origin (2006) Method for Qualification and Selection of Open Source Software (QSOS) version 1.6. Available at: http://www.qsos.org/download/qsos-1.6-en.pdf [Accessed: 15 March 2013].

Bowker, L. and Barlow, M. (2004) Bilingual concordancers and translation memories: A comparative evaluation. In Proceedings of the 20th International Conference of Computational Linguistics COLING-2004. Presented at the Second International Workshop on Language Resources for Translation Work, Research & Training, Geneva, Switzerland, 70-83.

BRR (2005) BRR Whitepaper 2005 RFC 1.

Calzolari, N., McNaught, J., Palmer, M. and Zampolli, A. (2003) ISLE Final Report. ISLE Deliverable D14.2. ISLE. Available at: http://www.ilc.cnr.it/EAGLES96/isle/ISLE_D14.2.zip [Accessed: 5 April 2015].

Cerezo, L. (2003) Hacia la evaluación de dos sistemas comerciales de memorias de traducción. In Entornos informáticos de la traducción profesional: las memorias de traducción. Granada: Editorial Atrio, 193-213.

Deprez, J.-C. (2009) QualOSS Assessment Methodology Version 1.1. QUALOSS Consortium. Available at: http://www.qualoss.org/deliverables/D4.5_StdQualOSSAssessmentMethod-v1.1.tar.bz2 [Accessed: 15 March 2013].

Deprez, J.-C. and Alexandre, S. (2008) Comparing Assessment Methodologies for Free/Open Source Software: OpenBRR and QSOS. In A. Jedlitschka and O. Salo (eds.) Product-Focused Software Process Improvement (Vol. 5089). Springer Berlin / Heidelberg, 189-203.

EAGLES. (1996) Evaluation of Translators' Aids. Available at: http://www.issco.unige.ch/en/research/projects/ewg95//node140.html [Accessed: 5 April 2015].

Eisele, A., Federmann, C. and Hodson, J. (2009) Towards an effective toolkit for translators. In Proceedings of the ASLIB International Conference Translating and the Computer 31. London: ASLIB. Available at: http://www.dfki.de/lt/publication_show.php?id=4586 [Accessed: 5 April 2015].

Filatova, I. (2010) Evaluación de herramientas y recursos informáticos (TAO y ofimática) para la traducción profesional: hacia la configuración de un entorno óptimo de trabajo para el traductor autónomo (doctoral thesis). Universidad de Málaga.

Flórez, S. (2012a) FOSS4Trans. Available at: http://traduccionymundolibre.com/wiki [Accessed: 25 July 2015].

Flórez, S. (2012b) OmegaT Results. Available at: http://traduccionymundolibre.com/wiki/OmegaT-Results [Accessed: 25 July 2015].

Flórez, S. (2013) Tecnologías libres para la traducción y su evaluación (doctoral thesis). Universitat Jaume I.

Flórez, S. and Alcina, A. (2011a) Catálogo de software libre para la traducción. Tradumàtica 9 (Software lliure i traducció), 57-73. Available at: http://revistes.uab.cat/tradumatica/article/download/5/6 [Accessed: 5 April 2015].

Flórez, S. and Alcina, A. (2011b) Free/Open-Source Software for the Translation Classroom: A Catalogue of Available Tools. The Interpreter and Translator Trainer (ITT): Volume 5, Number 2, 325-57.

García González, M. (2008) Free software for translators: is the market ready for a change? In Diaz Fouces, O. and García Gonzáliz, M. (eds.) Traducir (con) software libre. Granada: Comares, 9-31.

Gasser, L., Scacchi, W., Ripoche, G. and Penne, B. (2003) Understanding Continuous Design in F/OSS Projects. Presented at the 16th Intern. Conf. Software & Systems Engineering and their Applications, Paris. Available at: http://www.ics.uci.edu/%7Ewscacchi/Papers/New/ICSSEA03.pdf [Accessed: 5 April 2015].

Gow, F. (2003) Metrics for Evaluating Translation Memory Software (master's degree dissertation). University of Ottawa. Available at: https://www.ruor.uottawa.ca/handle/10393/26375 [Accessed: 5 April 2015].

Groven, A.-K., Haaland, K., Glott, R., Tannenberg, A. and Darbousset-Chong, X. (2011) Quality Assessment of FOSS. In INF5780 H2011: Open Source, Open Collaboration and Innovation, 73-91. Available at: http://publications.nr.no/directdownload/ publications.nr.no/Compendium-INF5780H11.pdf [Accessed: 5 April 2015].

Guillardeau, S. (2009) Freie Translation Memory Systeme für die Übersetzungspraxis. Universität Wien. Available at: http://othes.univie.ac.at/6863/ [Accessed: 5 April 2015].

Höge, M. (2002) Towards a Framework for the Evaluation of Translators' Aids Systems (doctoral thesis). University of Helsinki, Finland. Available at:

http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.18.3520&rep=rep1&type=pdf [Accessed: 5 April 2015].

ISO/IEC 9126 (2001) Software engineering. Product quality.

Kan, S. H. (2002) Metrics and Models in Software Quality Engineering (2nd ed.). Reading, Mass.: Addison-Wesley.

Lagoudaki, E. (2007) Translators evaluate TM systems – a survey. Multilingual, (March), 57-59.

Lagoudaki, E. (2008) Expanding the Possibilities of Translation Memory Systems: From the Translator's Wishlist to the Developers Design (doctoral thesis). Imperial College London.

Maslanko, K. (2004) A Comparative Study of Terminology Management Tools in Machine-Assisted Human Translation. Available at: http://www.transsoft.seo.pl/en/translator_tools.html [Accessed: 5 April 2015].

Quah, C. K. (2006) Translation and Technology. New York: Palgrave Macmillan Ltd.

Rico, C. (2001) Reproducible models for CAT tools evaluation: A user-oriented perspective. In Translating and the Computer 23. Presented at Aslib, London. Available at: http://www.mt-archive.info/Aslib-2001-Rico.pdf [Accessed: 5 April 2015].

Samoladas, I., Gousios, G., Spinellis, D. and Stamelos, I. (2008) The SQO-OSS Quality Model: Measurement Based Open Source Software Evaluation. In E. Damiani and G. Succi (eds.) Open Source Development, Communities and Quality — OSS 2008: 4th International Conference on Open Source Systems. Boston: Springer, 237–248. Doi:10.1007/978-0-387-09684-1_19.

Schulmeyer, G. G. (ed.) (2006) Handbook of Software Quality Assurance (4a ed.). Boston, London: Artech House.

TEMAA (n. d.) TEMAA Final Report. Available at: http://cst.dk/temaa/D16/d16exp-Contents.html [Accessed: 5 April 2015].

UNE-ISO/IEC 14598 (1998) Information Technology. Software Product Evaluation.

Wasserman, A., Murugan, P. and Chan, C. (2006) The Business Readiness Rating Model: an Evaluation Framework for Open Source.

Wiechmann, D. and Fuhs, S. (2006) Corpus linguistics resources: Concordancing software.

Wittmann, M. and Nambakam, R. (2010) OMM: CMM-like model for OSS. Qualipso Project.

Zerfaß, A. (2002) Evaluating Translation Memory Systems. Language Resources for Translation Work and Research, 49.

Usability of Free and Open-Source Tools for Translator Training Omegat and Bitext2tmx

María Teresa Veiga Díaz, Marta García González
University of Vigo, Spain

1 Introduction

In Spanish universities, free and open-source software (FOSS) is widely used in technical areas because of its usability, adaptability and low cost. Conversely, the use of these tools in the field of translator training has been minimal despite the existence of suitable software specifically developed for translation activities, such as OmegaT, Anaphraseus, bitext2tmx, Sun Open Language Tools, ForeignDesk or Transolution. In this context, GETLT was created to promote the use of FOSS both in translator training and professional translation, and to acknowledge the effort made by FOSS localization teams. After a short overview of the phases and results of research project PGIDIT07PX1B302200PR, Creación dunha plataforma docente GNU/Linux para a formación de tradutores – localizadores de software – subtituladores, funded by the Galician Government, within the framework of programme Incite, this chapter describes a particular research effort focused on testing the usability and applicability to translation training of free and open-source translation memory managers and text aligners with different texts types and genres.

1.1 The Background Project

The purpose of the initial project was to develop a computer environment for the training of translators and interpreters based on free, open-source software, more particularly a GNU/Linux distribution in a live-CD that could also be installed on the computer's hard disk, to be freely used at translation training university centers worldwide, and adapted to meet the particular needs of the educational programs at each university. The underlying idea was to develop an environment that could be used for translator training in all the different courses comprising a degree in translation and interpreting. It should facilitate the use of CAT tools for translator training, by removing the high costs of proprietary licenses, and also encourage the use of free, open-source software among students, future professional translators, thus covering the existing gaps within this group as concerns free software (Fernández García 2006a: 76-80; García González 2008: 9-31).

The activities in the project were arranged in four different phases, some of which were developed simultaneously rather than on a strict consecutive

basis. Although it is beyond the scope of this paper to discuss in detail each phase and the project's results (García González, 2013), a short description of the activities and main results follows:

Phase 1: Analysing training requirements in the different varieties of language mediation, by means of interviews to teachers and translation professionals, and **choosing a series of free software applications** running over GNU/Linux O.S. that were able to meet such requirements. The interviews were carried out both in situ and via e-mail and the information compiled was used as a basis for the subsequent phases of the project.

Phase 2: Following the above data (requirements and chosen applications), **generating a GNU/Linux distribution that was both live executable from a live-CD and installable on the computer's hard disk, targeted to the training of language mediation professionals**. The distribution was generated, based on Linux Mint Distribution, under the name of MinTrad (for a detailed description of this and other Linux Distributions for translators, see Sandrini in this same volume).

Phase 3: Disseminating the project's results within the university community: Results were presented at several conferences and also described in different papers and chapters during and after the duration of the project.

Phase 4: Documenting the distribution in a complete and sufficient manner, by preparing a comprehensive user guide for all the tools and applications comprising the distribution, and **testing the environment** both with students and with professional translators. This phase was planned as a long-term activity, as it could not be fully covered within the duration of the project. A short part of the testing effort is described in this chapter.

1.2   Documenting and Testing MinTrad

The distribution prepared under phase 2 of the project, MinTrad, included 30 computer-aided translation applications, among which one text aligner (bitext2tmx), and four translation memory managers (OmegaT, Anaphraseus, Transolution XLIFF editor, Sun Open Language Tool). As already mentioned, in addition to the preparation of the distribution, the project envisaged a phase focused on documenting and testing the applications in terms of their usability both in different types of translation courses and in professional translation situations. Here, usability is understood as the effectiveness, efficiency and satisfaction with which translation trainees and professionals achieve specified translation goals in a formative or professional environment, which is

in agreement with the standard definition of usability (ISO 9241). The satisfaction of translation trainees with the MinTrad distribution was preliminarily measured in previous phases of the project (García González, 2013) through a survey conducted among translation students. The survey included questions on their familiarity with FOSS, the complexity of the distribution, and the usefulness of MinTrad in translator training environments. Overall, fourth-year students, who had the opportunity to test the distribution with different types of texts, showed satisfied with the usefulness of the distribution in didactic settings and considered that it would be even more useful for use in professional environments. Yet, the survey did not include questions on the usability of specific tools or on the effectiveness and efficiency of the distribution. Accordingly, to complete the results of the previous phases of our research, the usability and applicability of two free and open-source computer-assisted translation tools included in the MinTrad distribution, namely OmegaT and bitext2tmx, were tested. The main purposes of the tests were (i) to determine the advantages and drawbacks of the tested applications as compared to similar proprietary software applications; and (ii) to determine the applicability of the translation memories generated by using the tested applications with different types of texts in the specialized translation classroom.

2 Materials and Methods

2.1 Tools

Two software applications were tested, bitext2tmx text aligner v. 1.0MO and OmegaT translation memory manager, versions OmegaT_2-2-2_04_Beta and OmegaT2.1.7_02 for Linux. As mentioned in section 1, both applications are free and open-source and are included in the MinTrad distribution. Bitext2tmx and OmegaT were tested under three operating systems, Windows XP, Linux MinTrad and MacOS X, insofar as it was assumed that the possibility of using the applications regardless of the operating system used was a big asset for translator trainees, who are not constrained to use a specific system. Actually, the computers available to our students both in free-access rooms and in classrooms have two partitions, one for Windows and another one for Linux.

Bitext2tmx   (http://bitext2tmx.sourceforge.net/doc/guide/en/Bitext2tmx.html) was originally developed by members of the Transducens research group at the department of languages and computer systems of the University of Alicante, Spain. As a text aligner, bitext2tmx allows for the creation of translation memories in TMX format by aligning an original text and its translation, both in plain-text format. The generated memories can be edited and aligned to provide better matches when used with any translation memory manager.

The tested text aligner was not further developed, such that no more recent versions are available.

OmegaT (http://www.omegat.org/en/omegat.html) is probably the most widespread free cross-platform translation memory application and has been the focus of several papers in the past few years (Carretero 2010; García 2010; Prior 2010). It is intended for professional use and commonly used by translation students at the University of Vigo. Among its features are: fuzzy matching, simultaneous use of multiple translation memories, user glossaries with recognition of inflected forms, more than 30 file formats (including Microsoft Office 2007 and later, PDF, HTML and XHTML, ODF, PO, and IDML/TTX/XLIFF/TXML), spelling checker, compatibility with other translation memory applications and interface to Google Translate. It is under constant development and has gradually incorporated new features. The most recent stable version of the application is OmegaT 3.1.9.

2.2   Methods

To determine the usability of the selected tools, the three components of usability, namely effectiveness, efficiency and satisfaction (Jordan 1998) were explored. Effectiveness was understood as the accuracy and completeness with which translators can achieve the relevant goals, i.e. a satisfactory alignment of two parallel texts or a satisfactory translation with a high per-centage of matches; efficiency was understood as the resources expended in relation to accuracy and completeness in terms of time, money and know-ledge required to use the tool and, finally, satisfaction was understood in terms of the comfort and acceptability of the system to the users. The method used to test the usability of the applications and to determine the applicability of the generated translation memories was divided into three phases: i) text alignment and generation of translation memories; ii) application to translation projects and iii) application to learning environments. The effectiveness and efficiency of the tools were analyzed in all three phases, while comfort and acceptability were studied mainly in the first phase of the analysis according to the following four criteria: accessibility and installation, interoperability, functionality and interface.

i) Text Alignment and TM Generation:

In the first phase, the translation memories that would later be fed into the TM manager were generated with bitext2tmx. To this end, a parallel text corpus was compiled. Also, a monolingual corpus was compiled to later test the usability of the OmegaT TM manager through the simulation of a number of translation projects. Both corpora included three sub-corpora, a sub-corpus of legal texts, a sub-corpus of economic texts and a sub-corpus of scientific and

technical texts. The selected texts were saved in different file formats, namely *.doc, *.txt, *.odt, *.rtf, *.pdf and *.ppt, such that the usability of both tools could be tested. The scientific sub-corpus was composed of only three genres: scientific papers, patient information leaflets (PILs) and game console user guides. The scientific papers included in the corpus were originally written in Spanish and translated into English, and focused on farm pro-duction and classification. The genres covered by the economic sub-corpus included corporate reports, annual accounts, cost and financial accounting reports, SAP user instructions, and press releases, while the legal sub-corpus included testaments, articles of incorporation, agreements, legal forms and EU legislation. Contrarily to scientific papers, the legal and economic texts were in their most part originally written in English and translated into Spanish, except in the case of EU legislation, of which no reference was found to which was the original text of the pair.

In some cases, individual translation memories were created from each pair of texts, but in other cases, as with testaments or corporate documen-tation (annual reports or UE legislation), the individual translation memories were merged with the help of an OmegaT plug-in, TMX-Merger, a Java command-line script for merging two or more TMX files. A total of 114 pairs of texts of different lengths, ranging 75 to 15800 words were aligned. In this phase, the effectiveness of bitext2tmx was determined by defining the accuracy with which the selected pairs of texts were aligned, and efficiency was determined based on the resources needed to complete the task. As per satisfaction, four criteria were considered: accessibility and installation, inter-operability, functionality and interface.

ii) Application to Translation Projects:

The memories generated in the first phase of our research were fed into the projects. A total of 11 translation projects were created, three of which corre-sponded to scientific and technical texts, another three to economic texts, and the remaining seven to legal texts. From among the seven legal translation projects, five corresponded to texts extracted from the EUR-Lex database and were analyzed as a unit. As in the text alignment phase, the selected source texts had different lengths so that the performance of the tool could be studied separately. For all text types, the texts selected for validation were similar to those used in the specialized translation classroom. In this case, the effective-ness of OmegaT was determined based on the number of 100% and fuzzy matches, and efficiency was analyzed in terms of the time and effort required to achieve an accurate and complete translation using the TM fed into the project. As in the first phase, the satisfaction of users was determined based on accessibility and installation, interoperability, functionality and interface.

iii) Application to Learning Environments:

After the texts were aligned and the performance of the generated TM was tested in OmegaT, the last phase of the project consisted in testing the tools in a specialized translation classroom, particularly in a scientific and technical translation course of the fourth year of the Degree in Translation and Interpreting. Three translation projects were created, one for each of the selected genres, a specialized paper, a PIL and a game console user's manual. The purpose of the test was to try both tools with the most common types of texts in the classroom and assess their benefits and drawbacks for translation trainees. Thus, students would learn: (i) to determine when and with which resources it is effective and efficient to use CAT tools; (ii) to identify the factors that affect the quality of a translation performed with these tools; (iii) to assess the suitability of the machine translation solutions provided by the TM manager. The criteria used to assess the usability and applicability of the tools in this phase were the same as in the second phase of the project, but the formative nature of the translation projects was considered.

3 Results

In this section, we present the results for the three phases of the project. First, we provide an overall assessment of the performance of bitext2tmx and OmegaT (for a thorough discussion of the quality of the translation memory manager, please see Flórez & Alcina in this same volume). Then we focus on the results of the application of both tools to particular translation projects, both professional and formative, for the three types of texts considered, business, legal and scientific, and technical.

3.1   Overall Assessment

3.1.1   Bitext2tmx

The main benefits of the text aligner included in MinTrad are related with accessibility and ease of use and installation, whereas the main drawbacks are related with efficiency. Bitext2tmx is a free and open source text aligner that requires no installation. It runs on the three operating systems tested, Windows, Mac and Linux, and generates .tmx files that are compatible with other CAT tools, both free and proprietary.

In didactic settings, bitext2tmx is highly applicable, because it is intuitive and easy to use for beginners. In addition, it runs smoothly with short, edited texts and the results for these texts are good, which makes it particularly suitable for use during the first years of the degree, when students start using CAT tools and translating very simple texts.

Despite these benefits, bitext2tmx has a number of problems related to its functionality that make it less efficient for use among advanced users or with longer texts than similar proprietary tools. Particularly, the following drawbacks have been observed during our testing:

Although the application runs on the three operating systems, it does not recognize files with hidden extensions in Mac OS X. Moreover, only *.txt files can be aligned, such that other types of files must be converted before alignment, which requires spending more time and effort.

Bitext2tmx does not allow for saving partial alignments, which can be seriously inconvenient when working with long texts. In addition, changes are not saved in case of a shutdown of the application, such that the users need to start over again, thus losing efficiency. Furthermore, alignment of more than one pair of texts per project is not enabled. Therefore, users cannot generate a single translation memory (TM) for several texts and each generated translation memory corresponds to a single pair of texts, thus forcing the use of a TMX merger. In bitext2tmx, alignment rules do not seem to consider language pair specificity, such as the average sentence length or the presence of graphical accents, which requires pre- or post-editing by the user in order to obtain a reliable TM. Moreover, some symbols and signs, such as those for percentages, decimals, semi-colons, among others, are often misinterpreted as full stops, which seriously affects segmentation and, therefore, effectiveness.

Finally, the application is not as user-friendly as similar proprietary tools because the interface lacks some functionality such as keyboard shortcuts, the scroll function for the translated-text window, or mechanisms for simultaneous selection of several lines of text. Yet, the "split by line break" functionality partially improves segmentation, particularly for tables and figures.

The above assessment suggests that bitext2tmx is a simple tool that can be useful for students who are involved with the translation of short, simple texts, but not for professional translators who prioritize efficiency.

3.1.2   OmegaT

OmegaT is an easy-to-use-and-install tool that runs on the three operating systems, although it requires reading the manual for the creation of new projects. In addition, OmegaT does not support every file extension, *.txt, *.docx and *.odt files are supported, but *.doc files are not supported. Yet, the main drawbacks of this free and open source application are related to its functionality.

As regards segmentation, OmegaT segments into paragraphs, with no segment expansion or shrinking enabled on the interface. If sentence segmenta-

tion is preferred, the text must be pre-edited and rules must be setup in the main menu, in Options → Segmentation. In addition, the application does not correctly identify the matches with long paragraphs, such that both effectiveness and efficiency are affected.

With regard to terminological extraction features, the application enables the generation of glossaries, but glossary terms cannot be automatically extracted, such that terms must be manually added to the project glossary. In addition, the glossary is necessary to retrieve specific terms because the application does not find matches by term. Yet, generating glossaries in OmegaT is very simple, insofar as glossaries are lists of words separated by a tab. In didactic settings, this is an advantage insofar as it allows students to reuse the glossaries prepared for every course and feed them into any project. In contrast, TMs from other projects or translators that have been generated with tools different from OmegaT, such as the bitext2tmx aligner, can be used as ancilliary translation memories but not directly imported into the master translation memory of the project, project_save.tmx, unless merged through the TMXMerger java command-line script. Working with many ancilliary TMs may unnecessarily slow OmegaT down, thus reducing the efficiency of the tool. In addition, ancilliary translation memories are read by OmegaT but not corrected during the project, which reduces the efficacy of the tool. Therefore, merging the TMs from other translation or alignment projects with the master TM speeds up the process and makes it more reliable. Nevertheless, merging .tmx files with TMXMerger requires some level of programming and might be tricky for some students, particularly for those who do not have specific training.

Another problem related with TM creation is the fact that wrong translations are not deleted when corrected unless they are stored in the main TM, which can affect the accuracy with which the relevant task is performed. Other efficiency issues are related to the creation of labels; OmegaT inserts "fuzzy match" labels that are not automatically removed when the final files are generated, such that users must remove these labels every time that an insertion is confirmed or when the final file is generated.

It should be noted that as versions OmegaT 2-2-2_04_Beta and OmegaT 2.1.7_02 for Linux were used in the test, some of the drawbacks referred to above might have been already corrected in later versions. In addition, despite the drawbacks, which can be rather limitative to professional users, OmegaT has many benefits for use by students. First, OmegaT is a free and open source tool that runs on the three OS tested and is already installed in MinTrad. The application includes a complete and relatively simple user manual and a readily accessible quick start guide that is very useful for

students who are starting to become acquainted with the application. The translation process is simple and intuitive, in contrast to project creation, which requires reading the manual. Once the project is created, the application is easy to use and the interface is user-friendly: it enables keyboard shortcuts, which speeds up the process, and incorporates machine translation options (Google Translate, Apertium, Belazar). The possibility to search Google Translate can be useful sometimes, but it must be handled with care in didactic settings, in order to avoid random use of the option by students.

Also, OmegaT retrieves up to five matches, indicating percent match and origin, which is useful when different unmerged TMs are used. In addition, the application allows alternative use of various files within the same project. Finally, the application offers some utilities, such as a text aligner and a tmx merger. Yet, as explained above, using these utilities requires specific knowledge of java script, which makes it complex for inexpert students.

In the following sections, the results of the applicability of the generated TMs for the translation of each text type and genre are discussed.

3.2   Applicability to Translation Projects

According to the test results, the applicability of the text aligner and the generated TMs depends strongly on text type and genre.

3.2.1   Business Texts

- Financial reports: good results were obtained both with TM manager and aligner when translating reports from the same company for different years. Otherwise, results were poor except for audit reports.
- General meeting agenda: again, results were highly satisfactory when the TM manager was used for the translation of agendas from the same company. When translating texts from other companies, though, results were poor except for legal fragments connected to companies law.
- SAP training presentations: several problems were encountered during the alignment of the (ppt) presentation, mainly connected to the conversion of text for alignment. However, after editing, the TM proved rather effective with similar SAP Training Documents.

3.2.2   Legal Texts

- EUR-Lex legal texts: overall, the use of the translation memories resulting from alignment of EU legal texts proved highly effective for the translation not only of other EU texts but also of acts from the different Member States that were adapted to EU law.

- Articles of incorporation: as in the case of financial reports, results with articles of incorporation were satisfactory when translating document amendments but rather poor with texts from different companies, except for legal fragments connected to companies law.
- Service agreements: although results were excellent with short texts, particularly with agreement forms, longer texts produced fewer match retrievals, particularly in sections containing specifications, which decreases the effectiveness of the tool.

3.2.3   Scientific and Technical Texts

- Specialized scientific papers: overall, the applicability of the generated TMs to scientific papers is very limited. Actually, the TMs generated from the text pairs used to test the aligner were useful only for papers with a high percentage of complete paragraphs repeated from previous papers. Accordingly, the usability of the tested tools for this text genre is very poor.
- User manuals of simple electronic devices: in contrast to the results obtained for specialized papers, both bitext2tmx and OmegaT showed highly usable for the translation of user manuals of different versions of simple electronic devices, provided that the quality of the aligned texts was good.
- Product information leaflets (PILs): the applicability of the generated TMs was excellent, in terms of both effectiveness and efficiency. Some comfort issues were observed, but the overall performance of the tool with this type of texts was very good.

3.3   Applicability to Translator Training Environments

To test the applicability of the tools to formative translation projects, the students of the course in Scientific and Technical Translation at the University of Vigo were asked to create three translation projects in OmegaT using the TMs generated in the first phase of our research, as mentioned earlier in this paper. In this section, the results of the activity are discussed.

- *Specialized scientific papers*: The results for effectiveness were very poor for this genre because of the extremely low percentage of 100% or fuzzy matches obtained by students. Actually, the TMs generated from the text pairs used to test the aligner were almost useless for the translation project tested in the classroom because of the low percentage of complete paragraphs repeated from previous papers. The number of matches retrieved with the tool was so low that it was highly inefficient. Efficiency could increase if the terminological management utility was improved, particularly to guarantee terminological consistency among papers by the same

authors. In addition, the solutions provided by Google Translate in this case were almost useless. Yet, the activity helped students learn to handle machine translation with care because of the evidently poor automatic translations retrieved. Consequently, despite the poor results, this type of project is useful as a formative tool for students insofar as they learn through practice that the applicability of the generated TMs for scientific papers is very limited.

• *User manuals of simple electronic devices*: very good results in terms of effectiveness and efficiency were obtained with instructive texts that corresponded to user manuals of different versions of the same game console. Provided that the selected texts correspond to simple devices, which are usually short, this type of text is highly applicable in the translation classroom for students who are not well-acquainted with text aligners and CAT tools. Yet, the quality of the translations strongly depends on the quality of the aligned texts. Therefore, the quality of the aligned translated texts will determine the teacher decision on whether it is efficient to use a text aligner to generate a translation memory. Alternatively, a good translation memory can be generated from the translation of short texts that are revised and corrected in the classroom, instead of generating a memory from translations available from the internet, as was the case of one of the texts tested in this phase of the project (see Figure 1).

290	To remove your game, first turn the power off.	Para remover el juego, primero apague el sistema.
291	Push the DS Game Card into Slot 1 until it clicks.	Empuje la Tarjeta de Juego de Nintendo DS hacia adentro de l...
292	It will automatically eject partway out of the slot.	De forma automática la tarjeta será expulsada parcialmente d...
293	Using Game Boy Advance Game Paks 1.	Uso de Cartuchos de Juego de Game Boy Advance 1.
294	Make sure that the Nintendo DS is turned off.	Asegúrese de que el Nintendo DS esté apagado.
295	2.	2.
296	Insert the Game Boy Advance Game Pak into Slot 2 on the front...	Inserte el Cartucho de Juego de Game Boy Advance en la ranur...
297	Make sure it is fully inserted into Slot 2.	Asegúrese de que el cartucho esté insertado por completo en l...
298	The label should face towards the bottom of the DS.	La etiqueta debería mirar hacia la base del Nintendo DS.
299	(Illustration 7)	(Ilustración 7)
300	3.	3.
301	Turn the power ON.	Encienda el equipo.
302	The game title will appear.	El título del juego aparecerá.
303	Touch the game title with the stylus to start the game.	Toque el título del juego con el stylus para empezar el juego.
304	Refer to the instruction manual for the game you are playing fo...	Refiérese al folleto de instrucciones para el juego que esté jug...
305	4.	4.
306	To remove a Game Pak, first turn the power off.	Para remover el juego, primero apague el sistema.
307	Push the cartridge out of Slot 2 with your thumb When you are f...	Usando el dedo pulgar como se indica, empuje el cartucho hac...
308	Do not wrap the AC Adapter cord around the DS.	Al dejar de jugar siempre apague su sistema y desconecte el...
309	When not in use, close the DS to protect the screens from dust...	No envuelva el alambre del Adaptador de Corriente en el Ninte...
310	NOTE:	Cierre el DS cuando no se encuentre en uso para proteger a la...
311	The power will NOT automatically turn off when the system is cl...	NOTA:

Figure 1: Alignment of an original text and a poor translation that makes the use of bitext2tmx inefficient.

• Product Information Leaflets (PILs): the performance of the text aligner and the TM manager was good for this genre. The stability of the macro-structure and phraseology of this genre makes it suitable for testing both

effectiveness and efficiency. A single pair or texts was aligned by students and fed into the project as a *.tmx file. Then, students were asked to trans-late the PILs for other presentations of the same drug, commercialized in Great Britain and Ireland with different names. A total of three PILs were translated using OmegaT but the process could be successfully extended to the PILs of every presentation of the same product. The results were excellent, and a total of 266 exact matches were found, which accounts for over 95% of the text (see Figure 2).

Figure 2: Almost automatic translation of a PIL using OmegaT.

PILs are commonly used in general and scientific translation courses and provide translation teachers with an excellent opportunity to successfully use free and open source CAT tools in the classroom. One of the benefits of using this genre is that text alignment is highly effective because of the fixed macrostructure and the length of the texts involved, which render the translation of similar texts efficient and effective. Yet, some drawbacks related

to satisfaction were observed by students. First, the text aligner and the TM manager segmented texts differently, such that post-edition was required after translation to avoid the presence of untranslated segments or format issues.

Second, when segments were not identical, the application did not recognize identical matches for some portions of text, such that the suggestions made by the application were not correctly prioritized (see Figure 3) and the suggested partial match was poorer than other available partial matches.
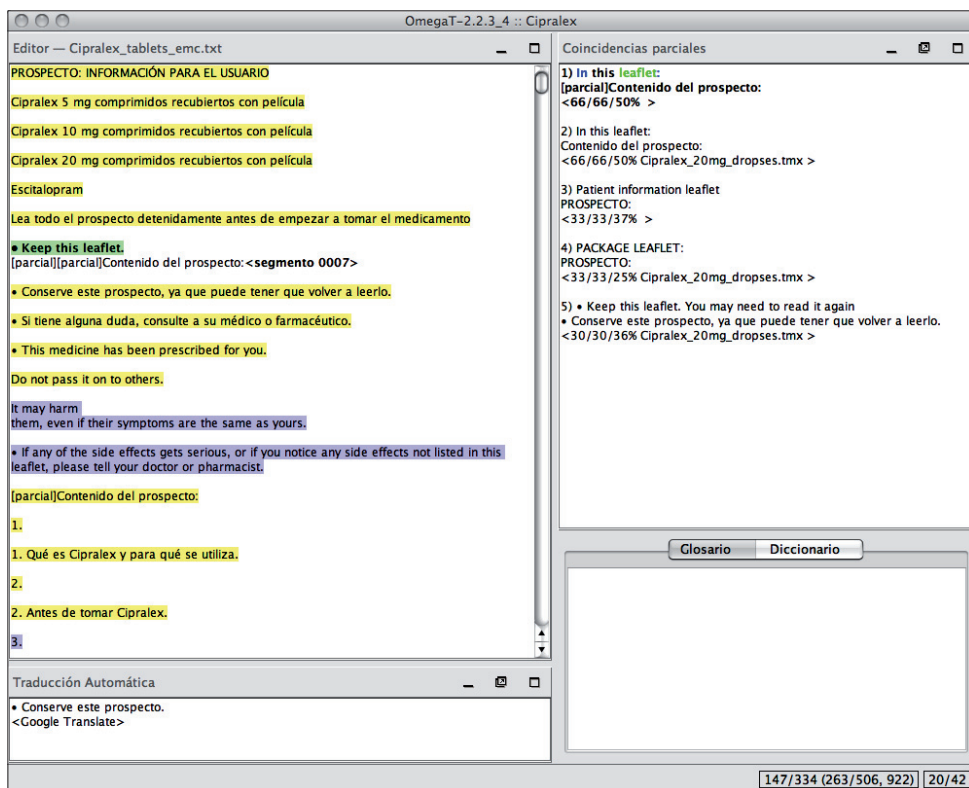
Figure 3: Wrong prioritazion of partial matches due to rigid segmentation rules.

4 Conclusions

As revealed by the results of the implemented translation projects, OmegaT performs much better than bitext2tmx in terms of effectiveness and efficiency, but the text aligner is easier to use, which increases the satisfaction of users. Overall, the usability of both bitext2tmx and OmegaT seems to be poorer than

the usability of similar proprietary software applications, but they can be used in translator training environments for a number of reasons.

First, bitext2tmx allows for the generation of TMs without the need to translate a large number of texts before generating a large TM that can be effective, thus reducing the time required to build useful translation memories from the texts translated in the classroom. Yet, there must be a balance between the time devoted to alignment and the time devoted to translation insofar as text alignment becomes inefficient if the percent of matches is low. Alternatively, students could use TMs available from the Internet. Yet, using this type of resources could be detrimental to students who are not well-acquainted with translation strategies.

Second, bitext2tmx helps students better understand how CAT tools work. When using an alignment tool first and then combining the resulting TM with a TM manager, students become aware of the manner in which texts are segmented and may check if this segmentation is appropriate for correct translation. This turns alignment into a relevant learning activity in the first phases of a translator training program.

Finally, OmegaT brings students closer to professional translation environments, in which productivity criteria prevail. On the other hand, using the tool with different types of texts enables them to determine its level of usefulness in different translation contexts. Particularly, they can realize that within the same course, a TM manager is highly productive for the translation of some genres and totally unproductive when translating other genres. Eventually, by using CAT tools and identifying their benefits and drawbacks, students realize that these tools are just tools, and not translators and that it is critical that they are competent translators before they can make the best of TM managers.

In sum, because the professional translation market increasingly demands the use of this type of tools, the translators-to-be need to have knowledge of the performance of the tools, not only of their benefits but also of their drawbacks. For this reason, bitext2tmx and OmegaT can be used as a "starter" in training students in the use of CAT tools despite the drawbacks observed during testing and reported in this paper.

References

Carretero, I. (2010) Free Software and Translation: OmegaT, a free software alternative for professional translation. In Boéri, J. and Maier, C. (eds.) Compromiso social y Traducción / Interpretación. Granada: ECOS.

Fernández García, J. R. (2006a) La traducción del software libre. Oportunidad de colaborar. LINUX Magazine, 19, 76-80. Available at: http://www.linux-magazine.es/issue/19/Educacion.pdf [Accessed: 28 January 2013].

García, E. (2010) Itzulpenak egiteko kode irekiko eta doako laguntzak. In Senez 43, 213-218.

García Gonzalez, M. (2008) Free Software for translators: is the market ready for a change? In Diaz-Fouces, O. and García Gonzaléz, M. (eds.) Traducir (con) Software libre. Alobolote (Granada): Editorial Comares, 3-31.

García González, M. (2013) Free and Open Source Software. In Translator Education. The MINTRAD Project. In The International Journal for Translation & Interpreting Research. 5-2., 125-148.

Jordan, P. (1998) An Introduction to Usability. London: Taylon & Francis.

Prior, M (2010) The open-source model. In ITI Bulletin, January-February.

Optimizing Process-Oriented Translator Training Using Freeware and FOSS Screen recording Applications

Erik Angelone
Kent State University, USA

1 Fundamentals of Process-oriented Translator Training:

1.1 Definitions, Models and Descriptions

As an empirically-drive pedagogical approach, process-oriented translator training, in a broad sense, focuses on enhancing learner awareness of how one translates. This overarching notion of 'how' can be approached from numerous, interrelated perspectives, including awareness of such phenomena as the nature of problems encountered and subsequent problem-solving tendencies (Angelone 2013a), segmenting behavior (Dragsted 2005; Hansen 2006), information retrieval tendencies (Alves and Liparini Campos 2009), general workflow patterns (Pym 2009), and cognitive ergonomics (Ehrensberger-Dow and Massey 2014). By deliberately shifting away from the translation product in and of itself as a relatively shallow snapshot of student performance, process-oriented training sets out to foster awareness of how this product was reached in the first place as a result of decision-making and strategy execution at the three fundamental loci of comprehension, transfer, and production. Given the fact that translation, at its very core, is a higher order cognitive task, process-oriented training approaches draw from numerous problem-solving models established within the cognitive process research community, such as that found in Figure 1:
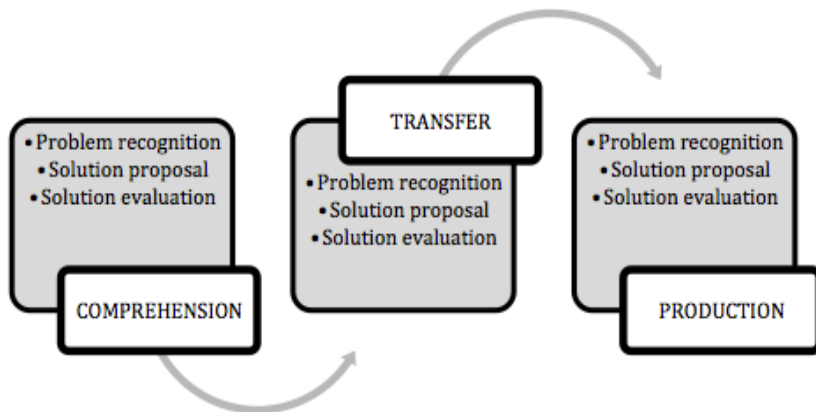
Figure 1: Loci and behaviors of problem-solving in translation (Angelone 2010).

Problem recognition involves knowledge assessment in relation to a given problematic aspect of the task at hand. There tends to be a breakdown in the natural flow of translation, with the most directly observable indicator thereof being an extended pause in translation activity. Solution proposal behavior involves strategy execution in response to the given problem, as indicated first and foremost by various forms of information retrieval. Whereas solution proposal concerns itself with generating options, solution evaluation involves narrowing them down in line with situational constraints. This is very much geared towards choosing among options, as driven by contextual factors and deliberate decision-making in light of them. All three of these behaviors (problem recognition, solution proposal, solution evaluation) can occur at any one the three loci (comprehension, transfer, production), often in a bundled, sequential fashion (Angelone and Shreve 2011: 120). Taken holistically, most directions in process-oriented translator training target some dimension of this particular problem-solving model.

1.2 Methods and Approaches

Process-oriented training began in earnest in the 1990s, when Kiraly (1995) called on trainers to shape a curriculum around optimal strategies, decisions, and behaviors exhibited by successful professional translators in authentic contexts. For the better part of that decade, translation process research and resultant pedagogical practices were driven by three primary methods: 1) Integrated Problem and Decision Reporting logs (Gile 2004), 2) think-aloud protocols (TAPs), and 3) keystroke logging. An IPDR log is a student-created running list of all problems encountered while translating along with correlating documentation of problem-solving strategies, rationales, and solutions used in addressing them. Creation requires students to temporarily break away from the translation task at hand to document content, which usually appears in tabular form in a separate document. IPDR logs are useful in generating whole-class discussion of problem-solving strategies in relation to a given text. However, the documented content is not always an entirely accurate reflection of the problems students faced, as revealed through mismatches between reported problems and actual errors that appear in corresponding translation products. This may by the result of still underdeveloped student self-reporting of problems, with problems tending to either go unnoticed or be defined in an incomprehensive fashion.

A think-aloud protocol consists of audio documentation of articulations representing thought processes that transpire over the course of translation. Students are instructed to engage in consistent, continuous verbalization in a relatively freeform manner. Retrospective analysis of recorded audio content can reveal problems and problem-solving tendencies in the form of extended

periods of silence, direct/indirect articulation, or a variety of speech dis-fluencies. Some students might feel uncomfortable with having to simul-taneously translate and articulate what is going through their minds, not to mention cognitively overtaxed by this dual task. As a result, it is advisable to keep the length of the texts to be translated short (200 words or less).

Towards the end of the 1990s, in response to documented shortcomings of translation logs and TAPs, keystroke logging became a methodology of choice for process-oriented training (cf. Hansen 2006). Here, a software application records all keystrokes, mouse clicks, deletions, and instances of cursor repositioning for purposes of retrospective analysis. Additionally, keystroke loggers document valuable temporal data, such as pause intervals and uninterrupted text segment durations, both windows to problems and problem-solving. The efficacy of keystroke logging as a lens to translation processes is evidenced by the fact that it is still very much a method of choice in the research community. Nevertheless, as depicted in Figure 2, from the student's perspective, making sense of highly granular data for purposes of self-reflection on problem-solving might be an onerous task.

Figure 2: Keystroke log output from Translog.

Over the past five or so years, a second generation of process-oriented translator training has come into existence, driven by two new methods on the cognitive process research front: 1) eye-tracking, and 2) screen recording. Eye-tracking technology, which documents visual attention data in the form of heat maps and gaze plots, holds great potential in helping trainers and trainees glean insight as to where students look on the screen and for how long when encountering and solving problems. To date, we have not seen much (if any) research on pedagogical applications of eye-tracking due to the high costs of existing commercial tools, but with the advent of open source eye-tracking applications, such as Opengazer (www.inference.phy.cam.ac.uk/opengazer), this may very well change in the near future.

Screen recording is made possible by a software application that captures all on-screen activity that occurs over the course of task completion, documenting such phenomena as extended pauses, information retrieval

(triggers and types of resources utilized), the textual level of target text gene-
ration, and revision tendencies. As is the case with TAPs, keystroke log out-
put, and the visual attention data made available through eye-tracking, when
using screen recordings, reflection on various aspects of the translation task
takes place during a retrospective session. Unlike eye-tracking, screen recor-
ding has gained firm footing in recent years as an optimal tool for process-
oriented training, particularly with the advent of freeware and open source
options. Reasons for this trend will be outlined in the next section. Table 1
below provides an overview of some of the advantages and disadvantages
associated with the five process-oriented training methods discussed in this
section.

Table 1. Process-oriented training methods

Method	Major advantage	Major disadvantage
Integrated Problem and Decision Reporting logs (Gile 2004)	Personalized content/ User-friendliness	Mismatch between perceived problems documented and actual errors in the product
Think-aloud protocols (TAPs)	Heightened cognitive focus when translating short (<200-word) texts	Cognitive/physical exhaustion from having to articulate all thought processes
Keystroke logging	Temporal data provides clear insight into allocation of cognitive effort	Students may lose sight of the "bigger picture" due to very granular data
Eye-tracking	Multiple triangulated levels of visual attention data (heat maps, gaze plots, saccades) = precision	Lack of portability/universality/ ecological validity
Screen recording	Highly visual rendition of problem solving results in heightened awareness of problems. (Angelone 2013a)	Thoughts underlying activity not always discernable = need for immediate retrospection post task completion

2 Screen Recording as a Preferred Tool

There are a number of reasons why trainers might want to turn to screen re-
cording as an optimal tool for freeware and FOSS-driven process-oriented
training. Firstly, recent empirical research has suggested that screen record-
ing, when compared with IPDR logs and TAPs as diagnostic protocols for

documenting student translation performance, is more efficacious in the domains of problem awareness and error mitigation (Angelone 2013a; Shreve et al. 2014). In a series of studies, students created logs, TAPs, and screen recordings in conjunction with various translation tasks and were asked to utilize the respective process protocol as a diagnostic tool of sorts to make any necessary changes to the corresponding translation products. When screen recordings were utilized for this purpose, fewer errors ultimately remained in the revised texts for the vast majority of students than when the other protocol types were used. This held true in tasks involving both self-revision and other-revision. The highly visual medium and manner of reflection would seem to potentially make problems more salient. This particularly holds true in light of the fact that students watch their performance as it originally unfolded in a very natural context. As previously mentioned, they do not have to do anything they would not otherwise already be doing while translating besides pressing record and stop. They do not have to work in an otherwise foreign interface. They do not have to make sense of numbers generated by an overly complicated analytic software application. They can engage in analysis from the comfort of their own homes on their own computers, thanks to cross-platform options. At the click of a mouse, they can fast-forward, rewind, and pause videos so that analysis transpires at their own preferred pace in a learner-centered fashion that is much less dependent on the trainer.

When screen recording technology was first integrated for research and training purposes, options were somewhat limited, with the vast majority of initiatives relying on Camtasia Studio, a proprietary application launched by the company TechSmith in 2002. At the time of writing, a single user license at education pricing rates costs $179 USD. Over the past decade, freeware and open source alternative options have entered the scene, as outlined below in Tables 2 and 3. Screen recording has evolved to become truly universal in the sense that it is not restricted to any one operating system/platform, output format, or programming language. Trainers and trainees should be able to find an application that best meets their potentially unique needs and preferences in terms of technical requirements and features. It is important to note that the FOSS (free and open source) options offer more or less the same level of functionality and range of features as their commercial counterparts. Quality is in no way sacrificed.

2.1 Screen Recording Features from a Training Perspective

Tables two and three below provide information on a selection of free screen recording applications, with variation at the levels of classification (freeware, freemium, or open source) and operating system (Windows, Mac OS, or Linux). These six applications, rather than representing an exhaustive list of

all that is available, were selected for inclusion based on a level of functionality and range of features that compare with Camtasia Studio as a commercial application benchmark. A brief overview of the various features with an eye towards pedagogical applications in the context of process-oriented translator training will be followed by descriptions of concrete learning activities.

Audio Recording (AUDREC)

This feature enables translators to capture audio documentation of their problems, problem-solving strategies, and general thought processes in the form of recorded articulations. The obtained audio data, in essence a TAP, parallels visual data representing on-screen activity, thereby providing a more granular depiction of translation processes. From the perspective of problem awareness training, students could be encouraged to focus in on such things as direct/indirect articulation of problems, extended periods of silence in articulation, and various speech disfluencies in retrospective analysis of their work.

Webcam Recording (WEBCAM)

With this feature, translators and translator trainers can obtain documentation of things like facial expressions, body language, and physical reactions in a broad sense in conjunction onscreen activity. In this sense, webcam data can be regarded as the non-verbal counterpart to the verbal data captured through audio recordings, adding another layer of granularity to the documentation and subsequent analysis of translation processes.

Scheduled Recording (SCHED)

This feature provides the option of starting and stopping recording at pre-set times and for a pre-set duration. If, for example, students or trainers want to examine how translation processes vary at different points of the task as it progresses (i.e., what do students do for the first ten minutes or the last ten minutes,), this feature could provide such snapshots for retrospective analysis. Obtaining such snapshots might also be interesting in documenting translator style and how this style might vary in situations involving timed vs. untimed tasks.

Real-time Pausing (PAUSE)

With this feature, translators can pause recording and continue at a later time, implying that there wouldn't be a need to complete the entire translation task in one sitting. This becomes particularly helpful in the context of lengthy texts, where the translator would likely be more inclined to take frequent breaks. This feature would also be helpful in situations where the trainer or trainee is

looking for documentation of only a specific aspect of the translation task, such as information retrieval tendencies. Everything else could be filtered out of the screen recording using real-time pausing.

Post-editing (EDIT)

This feature enables cutting, merging, or adding frames within a given screen recording after it has been created. This gives the trainer the option of creating montages to highlight such things as different ways of approaching the same problematic text passage or the execution of the same particularly efficacious problem-solving strategy at different locations in the task.

Annotation (ANNOT)

The annotation feature gives students and trainers the option of inserting various comments, such as documentation of observations, explanations underlying various strategies, etc., directly into the created screen recording. Depending on the application being used, the annotation may take the form of text, graphics, or even embedded videos.

URL-based Sharing (SHARE)

Screen recordings, particularly those representing longer translations (upwards of an hour), can be quite large in terms of file size, making sharing via email or e-learning platforms potentially problematic. The screencast sharing feature basically stores the recordings in an on-line repository that can then be accessed by others via designated urls. This is a nice way of sharing files based on permission settings and overcomes space limitations associated with other options.

Unlimited Recording Length (LNGTH)

Some screen recording applications have a set maximum recording time before automatically shutting off. Others enable recording videos of unlimited length, implying fewer restrictions on variables such as text length and difficulty, not to mention one less thing for trainers or trainees to worry about in an attempt to preserve ecological validity.

Table 2. A selection of screen recording options

Application	Publisher	Source	Classification	OS
Blueberry Flashback Express	Blueberry Software	http://www.bbsoftware.co.uk/	Freeware	Windows
CamStudio	CamStudio	http://camstudio.org/	Open Source	Windows
EZVid	EZVid, Inc.	http://www.ezvid.com/	Freeware	Windows
Open Broadcaster Software	obs.jim@gmail.com	https://obsproject.com/	Open Source	Windows Mac OS Linux
Quicktime	Apple Inc.	https://www.apple.com/quicktime/download/	Freemium	Windows Mac OS
recordMydesktop	Martin Nordholts	http://recordmydesktop.sourceforge.net/	Open Source	Linux

Table 3. A comparison of applications by features

Application	AUDREC	WEBCAM	SCHED	PAUSE	EDIT	ANNOT	SHARE	LNGTH
Blueberry Flashback Express	+	+	+	-	+	-	+	+
CamStudio	+	+	-	+	-	+	-	+
EZVid	+	+	-	+	+	+	+	- (45 min)
Open Broadcaster Software	+	+	-	+	+	+	+	+
Quicktime	+	-	-	-	+	-	-	+
recordMy desktop	+	-	-	+	-	-	-	+

3 Pedagogical Approaches and Learning Activities Using Screen Recording

Given the constellation of features outlined above, screen recording has proven to be a versatile application for purposes of process-oriented translator training. This section will describe a series of screen recording-based learning and assessment activities to facilitate learning along these lines.

3.1 Self-awareness of Problems

As mentioned above, empirical research on student problem-solving has indicated a tendency for problems to often go unnoticed (Göpferich 2009). Furthermore, what students assume to be problematic often represents only a narrow scope of what is truly problematic from the perspective of errors that

result in their translations. Having learners create screen recordings in conjunction with their translations establishes empirical grounds for diagnostic self-reflection and a mechanism for training problem awareness at a much more granular level than possible when examining the product alone. Prior to having students engage in self-reflection, it is paramount for trainers to guide them through the process and introduce various focal points, starting with potential problem indicators embedded in the screen recordings. Primary problem indicators include extended pauses in screen activity, instances of information retrieval, and revisions, among others. When analyzed empirically by students on a regular basis and across a variety of translation tasks, these are the kinds of phenomena that can yield a more holistic understanding of the nature of problems and problem-solving.

If students have the opportunity to submit drafts of a given translation, analysis of screen recordings in this capacity can serve as an important error detection editing stage prior to re-submission. Students could also be asked to write up a reflection on their problems and problem-solving tendencies using the following questions as prompts: 1) What tended to pose problems based on observed occurrences of extended pauses in screen activity? 2) How would you describe the nature of these problems from the perspectives of textual level (lexis, syntax, stylistic) and locus (comprehension, transfer, production)? 3) Which resources did you tend to utilize in addressing the problems and why? 4) In retrospect, was there anything that surprised you about the problems you encountered and the manner in which you went about solving them? 5) In retrospect, would you have done anything differently? Why? The documentation of these observations could serve as formal assignments or as a springboard for in-class discussion during workshopping sessions. Given the annotation feature described above, observations could be documented in the screen recording environment itself, eliminating the need for a different (separate) application for this purpose. Assignments could be submitted using the url-based sharing application inherent to many screen recording tools. Free and open source applications, in particular, have greatly advanced this 'all-in-one' approach, where student and instructor comments can be directly embedded in screen recordings, making file management and transfer that much easier.

3.2 Re-tracing Errors in the Product through the Processes

When it comes to feedback on their performance, students often have little more than marked-up errors in their translations to go on. These markings likely provide them with quantitative insight regarding the types of errors they make, yet often shed no light on why these errors may have occurred in the first place from a process-oriented perspective. For example, an error code

might reveal to the student that a terminology error has occurred, but he or she might not know why. Was it a result of inaccurate information retrieval? Was it a result of simply not knowing what the term means? Did he or she have the right term first and then go back and erroneously change it during a revision stage? Was the term's usage cross-checked using parallel texts? Did the terminological error co-occur with extended pauses to signal a potential problem? Screen recording documentation would enable the student to re-trace the error and answer these questions in obtaining a clearer insight into its nature, transcending beyond the textual level alone, as indicated in the mark-up. As a very basic learning activity, students could be asked to re-trace all of the errors in their translations and comment on why the errors may have occurred based on what they observe in their screen recordings. This form of self-assessment adds a much-needed procedural dimension to helping students understand the nature of errors.

3.3 Watching and Learning from Virtual Professionals

Screen recording can also be an effective way to introduce students to the problem-solving tendencies of professional translators. This can best be accomplished by having professionals create screen recordings while trans-lating the very same texts that students will be asked to translate, establishing grounds for comparative process analysis (Angelone 2013b). Students could be asked to focus on similarities and differences, at a very basic level, thereby enhancing awareness of multiple problem-solving pathways. Trainers could use this comparative approach as a way of modeling best practices from an expertise perspective, where students are asked to comment on the behav-iors and strategies of particularly successful professionals. Additionally, stu-dents could be asked to comment on where the professionals seem to struggle, or where their own problem-solving approaches could be regarded as more efficacious than those of the professionals. This latter activity can be particularly helpful in motivating learners and boosting their self-confidence. Additionally, it presents the real world of professional translation as being within reach.

3.4 Workshopping the Process

In a product-oriented training environment, a common pedagogical approach involves comparative analysis of translation products on a sentence-by-sentence basis. Screen recording enables an approach that focuses on how TT solutions were generated in the first place, also in a comparative fashion. Using the aforementioned editing feature, trainers can create collages representing multiple problem-solving approaches in conjunction with select text passages. Instead of reading multiple target text options on screen,

students would watch multiple target text options emerge in real time. This learning activity could be centered around an examination of what unfolds in conjunction with text passages that the trainer regards as being 'rich points' (PACTE 2011: 38), or predicted sources of disturbance (Hansen 2006). Alternatively, depending on how much lead time is available prior to in-class workshopping, the trainer could create collages based on observed, patterned problems. This would be conducive in situations where there is a potential mismatch between passages the trainer assumes will be problematic and passages that actually prove to be problematic based on evidence documented via screen recording.

3.5 Snapshots of Performance for Process-oriented Assessment

Formal assessment of translation using screen recording technology is a domain in which a vast amount of research is still waiting to be done. At the Zurich University of Applied Sciences, screen recording is being utilized in the context of assessing borderline entrance translation exams (Massey and Ehrensberger-Dow 2013). Given the fact that the translation product represents a somewhat limited view of student performance, taking a closer look at underlying processes might provide a more accurate (or at least more granular) reflection on student performance patterns (and potential) on the whole. That being said, given the length of screen recordings, holistic analysis of screen recordings in conjunction with each and every translation becomes less of an option for the individual trainer, particularly in the context of a higher enrollment class. To compensate for this, using the scheduled timer feature, trainers can utilize screen recording to capture a shorter representative sample of a larger translation task to analyze in conjunction with grading of the translation product. Quantitative metrics currently are not in place to guide process-oriented grading as such. In this case, ungraded feedback on processes can serve as an ideal complement to a concrete grade/letter score assigned to the product, even if based on only ten or so minutes of content.

4 New Horizons through a Freeware/FOSS Lens

Given the still predominately product-oriented focus of translator training and assessment (Dam-Jensen and Heine 2009: 1) and the fact that extensive feedback in the world of professional translation is seldom present, both students and professionals rely to a large extent on self-assessment in gauging their performance. In this sense, screen recording, as a process-oriented self-assessment tool, should be on equal footing with other freeware and FOSS applications constituting assistive translation workbenches, such as tuxtrans (Sandrini 2007) or CasMaCat (Koehn et al. 2012). The CasMaCat

open source workbench is already geared towards 'automatic analysis of translator behavior' (Alabau et al. 2013: 105) thanks to a logging and replay component based first and foremost on eye-tracking and keystroke logging technology. The inclusion of a screen recording component would likely enhance user-friendliness from the student's and trainer's (as opposed to the researcher's) perspectives in particular.

Interestingly, unlike what is the case for such CAT applications as translation memories and terminology management systems where industry-leading commercial options have emerged, there is no commercial screen recording benchmark against which FOSS and freeware options would need to compete. This gives each individual user (whether trainer, trainee, or professional) the freedom to pick and choose from a variety of screen recording options that best suit his or her unique needs and preferences without feeling forced into choosing a set industry standard and without having to worry about licensing or budgetary constraints.

In summary, as a CAT tool whose potential as a vehicle for enhancing process awareness is just now being realized in academic contexts, screen recording truly embraces portability, flexibility, and opportunities for customization envisaged by open source as a development model. It is hoped that the ideas presented in this paper will further motivate trainers, trainees, professional translators, and the language industry at large to explore how freeware and FOSS screen recording can be integrated to enhance translation pedagogy.

References

Alabau, V., Bonk, R., Buck, C., Carl, M., Casacuberta, F., García-Martínez, M., González, J., Koehn, P., Leiva, L., Mesa-Lao, B., Ortiz, D., Saint-Amand, H., Sanchis, G. and Tsoukala, C. (2013) CASMACAT: An Open Source Workbench for Advanced Computer Aided Translation. The Prague Bulletin of Mathematical Linguistics 100, 101-112.

Alves, F. and Liparini Campos, T. (2009) Translation Technology in Time: Investigating the Impact of Translation Memory Systems and Time Pressure on Types of Internal and External Support. In S. Göpferich, A.L. Jakobsen, and I. Mees (eds.) Behind the Mind. Methods, Models and Results in Translation Process Research. Copenhagen: Samfundslitteratur Press, 191-218.

Angelone, E. (2013a) The impact of process-protocol self-analysis on errors in the translation product. Translation and Interpreting Studies 8(2), 253-271.

Angelone, E. (2013b) Watching and Learning from 'Virtual Professionals': Utilizing Screen Recording in Process-Oriented Translator Training. In D. Kiraly, S. Hansen-Schirra and K. Maksymski (eds.) New Prospects and Perspectives for Educating Language Mediators. Tübingen: Günther Narr, 139-155.

Angelone, E. (2010) Uncertainty, uncertainty management and metacognitive problem solving in the translation task. In G. Shreve and E Angelone (eds.) Translation and Cognition. Amsterdam: John Benjamins, 17-40.

Angelone, E. and Shreve, G. (2011) Uncertainty Management, Metacognitive Bundling in Problem-Solving, and Translation Quality. In S. O'Brien (ed.) Cognitive Explorations of Translation. London/ New York: Continuum, 108-130.

Dam-Jensen, H. and Heine, C. (2009) Process Research Methods and Their Application in the Didactics of Text Production and Translation. trans-kom 2(1), 1-25.

Dragsted, B. (2005) Segmentation in translation – differences across levels of expertise and difficulty. Target 17(1), 49-70.

Ehrensberger-Dow, M. and Massey, G. (2014) Cognitive Ergonomic Issues in Professional Translation. In J. Schwieter and A. Ferriera (eds.) The Development of Translation Competence: Theories and Methodologies from Psycholinguistics and Cognitive Science. Newcastle upon Tyne: Cambridge Scholars Publishing, 58-86.

Gile, D. (2004) Integrated problem and decision reporting as a translator training tool. JoSTrans 2, 2-20.

Göpferich, S. (2009) Towards a model of translation competence and its acquisition: the longitudinal study TransComp. In S. Göpferich, A.L. Jakobsen, and I. Mees (eds.) Behind the mind: Methods, models and results in translation process research. Copenhagen: Samfundslitteratur Press, 11-37.

Hansen, G. (2006) Retrospection Methods in Translator Training and Translation Research. The Journal of Specialised Translation 5, 2-41.

Koehn, P., Alabau, V., Carl, M., Casacuberta, F., García-Martínez, M., González-Rubio, J., Keller, F., Ortiz-Martínez, D., Sanchis-Trilles, G. and Germann, U. (2012) CasMaCat: Cognitive Analysis and Statistical Methods for Advanced Computer Aided Translation. Available at: http://www.casmacat.eu [Accessed 27 June 2015].

Kiraly, D. (1995) Pathways to Translation: Pedagogy and Process. Kent, Ohio: Kent State University Press. Massey, G. and Ehrensberger-Dow, M. (2013) Evaluating translation processes: opportunities and challenges. In D. Kiraly, S. Hansen-Schirra, and K. Maksymski (eds.) New Prospects and Perspectives for Educating Language Mediators. Tübingen: Günther Narr, 157-180.

PACTE Group. (2011) PACTE Translation Competence Model: Translation Project and Dynamic Translation Index. In S. O'Brien (ed.) Cognitive Explorations of Translation. London/ New York: Continuum, 30-56.

Pym, A. (2009) Using Process Studies in Translator Training: Self-discovery through Lousy Experiments. In S. Göpferich, F. Alves, and I. Mees (eds.) Methodology, Technology and Innovation in Translation Process Research. Copenhagen: Sammfundslitteratur Press, 135-156.

Sandrini, P. (2007) *tuxtrans* GNU/Linux Desktop for Translators. Available at: http://www.uibk.ac.at/tuxtrans/ [Accessed 23 June 2015].

Shreve, G., Angelone, E. and Lacruz, I. (2014) Efficacy of Screen Recording in the Other-Revision of Translations: Episodic Memory and Event Models. MONTI 1, 225- 245.

Openness in Translators Training: a Case Study

Adrià Martín-Mor, Ramon Piqué i Huerta, Pilar Sánchez-Gijón
Universitat Autònoma de Barcelona, Spain

1 Introduction

It is very common for NGOs and public institutions to turn to translation training centres to have their main digital resources translated. Whether it is a website, internal documentation or even termbases, these institutions offer translation students a training opportunity to work with real products. But since translation training is, in the end, something more than just getting a particular text translated, the success of such training will depend on establishing an appropriate training context.

In the collaborative venture presented here, the interest by both parties came from another level from the outset. This was a collaborative venture between the Servei de Publicacions (SP – Publications Service) at the Universitat Autònoma de Barcelona and the Tradumàtica research group at the same university. The SP, which manages UAB publications, decided to introduce the OJS software package as a standard for managing and publishing academic journals. OJS is a free software for managing and publishing journals developed by the PKP consortium. This software has been developed by many within the international academic community and with a focus on localisation into various languages. One of the journals currently published through this system is Revista Tradumàtica, run by the research group of the same name. The Tradumàtica research group (www.tradumatica.net) is concerned with research into translation technologies in the broad sense, ranging from the description of the analysis of the translation process from the digital perspective to translator training in these specialised professions.

2 Choosing the Product

The collaborative venture between SP and the Tradumàtica research group to localise PKP software into Spanish and Catalan started during the academic year 2011-2012 and has been going on ever since. OJS caught the attention of the research group for a variety of reasons:

- Specific community asset transfer. Being able to make use of the interfaces and help files of the updated versions of PKP software in the most commonly-used languages at UAB (Catalan and Spanish) would clearly foster the use of this platform by editors and potential readership alike.

Therefore, it is an asset transfer towards the Spanish and Catalan speaking academic community.

- Localisation of FOSS software. It is important, when designing a collaborative localisation project involving students, to choose ethically correct proposals. In this respect, the localisation of PKP software means firstly promoting an initiative which facilitates free access to knowledge and, secondly, being FOSS software, its localisation does not involve students in any profit-making activity. Furthermore, as Diaz Fouces (2011: 10) puts it, "[l]a definición de un espacio profesional autónomo y digno supone no renunciar a mantener el mayor grado posible de control sobre los procesos de traducción" ("The definition of an autonomous and dignified professional space implies not waiving to keeping the highest possible control over the translation processes", our translation).

- Enhancing the product. PKP software (mainly OJS and OMP) is designed to manage and publish journals and monographs. Its development is supported by researchers involved in publications of an academic nature. Along these lines, all manner of editorial processes were envisaged during its development. Nonetheless, some design solutions adopted to facilitate the localisation of the software into other languages were not deemed the most appropriate by the Tradumàtica research group. On the basis of its experience, the group proposed software design enhancements aimed at overcoming these design problems.

- Being able to promote the use of minority/ised languages. Finally, localising into Catalan also involved standardisation. Although the main user community can work with the software directly in the Spanish or in the English versions, the localisation of the software into Catalan is currently possible within a context of standardised use due to efforts in recent years to standardise Catalan in the field of technology. Furthermore, by following the most widely used guidelines for localisation (for example, as regards the use of specialised vocabulary linked to software), we also collaborate in spreading its use among the community of users (Softcatalà 2010).

3 The Added Value of the Project

Once it was decided that PKP was an appealing initiative for the research group, the question of how collaboration could be established was posed. One of the most visible dimensions of the Tradumàtica research group is the Tradumàtica Masters. This is an M.A. programme oriented towards preparing

students for the professional world with company internships and an M.A. final project (TFM, from the Catalan 'Treball de Fi de Màster'), focused on mastering the translation process and the localisation of digital products. The M.A. coordinators decided to use OJS as a product which would be localised within the framework of the TFM. Students would thus be able to put into practice all the knowledge and competencies acquired during the M.A. programme through the management, translation and testing of the software, and at the same time reflect on the localisation process.

The proposal to localise PKP software within the framework of the M.A. offered advantages for the students well worth laying out. As regards our interests as a translation training centre, it offers the opportunity of providing students with real software and, at the same time, sufficient volume of work to justify all the localisation work carried out by the approximately thirty M.A. students. It allows us to manage the project through small work groups of between 3 and 5 students. For each brief, every two weeks, the students have to change task and adopt the role of project manager, translator, proof-reader and tester. As this is real software, their translation might be subject to all the conditioning factors of a real localisation project in terms of processes, phases, tools, problems, etc. Furthermore, software updates provide sufficient volume for the entire group. Therefore, introducing PKP software which stu-dents could localise as part of their training meant added value to their training and the M.A. programme. Its inclusion in the form of a TFM has proven to be a good move as well, since students are able to combine it with company internships, during which they are exposed to other products and workflows.

By localising real software under real professional practice conditions, the team of researchers/teachers involved in the project had the opportunity as well to delve further into the development of a project of this nature. Although as group researchers we are continually in touch with the professional trans-lation sector, our obligations as full-time lecturers at the UAB prevent us from being directly involved in projects such as this. Therefore, managing both the localisation project and the learning process of the students has been of major interest for the members of the teaching team involved. Real work with the most commonly-used tools, solving specific problems corresponding to phases of the process, etc., has meant total involvement by the teachers in managing and carrying out the localisation projects. For these reasons we believe that the work with PKP represents added value for the group's research members and consequently for the M.A., given that all this will be directly applied in future M.A. classes.

In fact, following the track of the most recent professional practices allows scholars to achieve two different objectives. Firstly, as translator trainers they have the chance to test new training models that guarantee students achieve the professional competencies needed in the translation industry. Secondly, researchers are able to take advantage of these training experiences and undertake studies to come to theoretical or empirical conclusions. Studies that measure the impact of professional practices in terms of quality or productivity are of special interest for the translation industry, but equally studies that shed light on theoretical or methodological issues of particular interest to the field of Translation Studies. This approach to Translation Studies research follows Munday's statement (2008: 179): "the emergence of new technologies has transformed translation practice and is now exerting an impact on research and, as a consequence, on the theorization of translation."

The accumulative experience gained by the Tradumàtica research group teachers from managing this localisation project has clearly allowed for developing the contents and competencies which they deal with in the M.A. in the direction of an entirely professional context. We have been able to develop our teaching models and allow more room for competencies such as teamwork and self-learning skills (regarding translation tools and problem solving). The teaching angle of this experience has allowed us to tackle competencies such as those mentioned above from a more genuine and professional perspective.

This experience has also allowed us to put into practice theoretical models developed by the group's researchers concerning the development of research projects. On the basis of this experience we have been able to develop these models according to changes in the translation profession which are becoming more and more important in the professional sector, such as machine translation and post-editing, or incorporating the specific quality control parameters required of international standards. This development from a theoretical slant has been one of the major benefits of the OJS localisation project for the Tradumàtica group.

As a consequence of the evolution of theoretical models, this project has also allowed the researchers to identify new research areas of use to society. One of the aims of the entire research group is that its research implies a return for society. Sometimes, it is difficult to measure this return. Other times, this return is too specific, and it ends up becoming a transfer of assets between universities or research centres and particular sectors of society. In fact, the majority of calls for research projects nowadays are aimed at facilitating research that offers a return for society and which contributes to the economic, productive, social and cultural development of the community.

Following this line of reasoning, it should be pointed out that participating in projects such as the PKP software localisation allows researchers to identify much more precisely the objects of study upon which public research can have an impact and which could result in a greater return for the community. A specific case in point is our community, in which we have a professional translation sector comprised of many small companies, in many cases one-person businesses, and a significant fabric of medium-sized companies employing up to 20 staff. By identifying these research objects whose development can benefit professionals in the translation market – and, indirectly, any professional sector –, the return of our work as researchers to society is guaranteed.
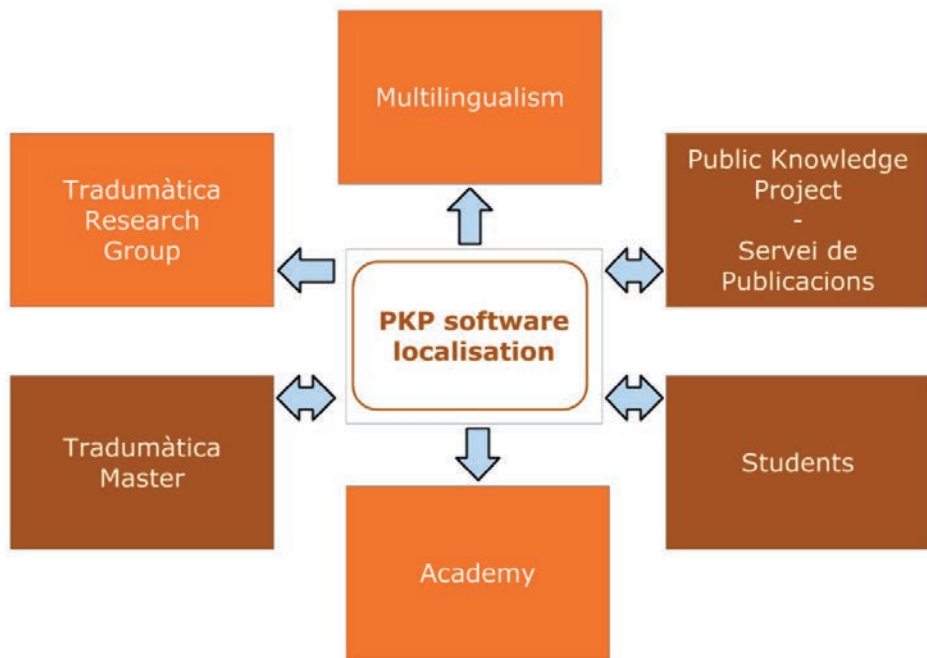
Figure 1: A multifaceted approach to PKP localisation.

4 The Key to Success

Despite all the advantages mentioned earlier, it also must be mentioned that the development of projects such as this are very demanding on all those involved. On the one hand, the NGO which provides the software to be localised has to act as a client in all senses. In our case, the SP at UAB has

to take on the responsibility of preparing all the files to be localised and gives an introductory training session to the software for the students involved in the localisation of the program. More importantly, they succeed in the challenge of having to resolve terminology and language use doubts within time frames of less than 24 hours, in order to guarantee that these doubts do not become an obstacle to meeting the deadlines set for each translation brief. They even developed a tool to facilitate real testing of the software before the localisation project was finished.

From the management point of view, without a doubt one of the keys of the success of this project is that everyone is able to collaborate via a server (groups of students, coordinators and terminologists), in such a way that the resources used (essentially translation memories and terminology databases) are queried and edited simultaneously by all participants. This eases speeding up processes within each group and thus bringing forward deadlines. On the other hand, however, this requires investing time and effort in managing the task prior to the translation brief.

For the teachers/researchers who took part in this project, this requires maximum commitment. Given that they assume two roles – teachers guiding the learning process and managers of the global project –, they have to be very flexible and accommodate deadlines to the development of the project itself. By acting as managers who commission specific translations with deadlines for each work group, the turnaround time for answering queries and solving problems has to be very short. This means that the teachers must have round-the-clock access to the resources used to develop the project: tools, materials, agendas, calendars, etc., and update, modify or adapt them to whatever situation that might crop up. In addition, by also managing the learning process, they have to provide themselves with the appropriate space so that students can get to the right conclusion for each problem they encounter, guaranteeing optimal results for the training of the students. This dual role demands a high level of commitment to the project not only while it is underway but also during the preparatory and concluding phases.

5 Dealing with Quality

The PKP localisation project to Catalan and Spanish may be seen at the intersection between a crowdsourced translation, a professional project and a students' assignment. Despite this idiosyncratic nature, different actions were carried out in order to ensure the quality of the final product, even if – as stated above – localising a real product increases per se the students' awareness of the importance of quality (the students were informed beforehand that

their names will appear in the contributors section of the PKP wiki at https://pkp.sfu.ca/wiki/index.php?title=Translating_OxS).

First of all, after the translators' final checks, each group carried out a crossed revision of the files translated by their own translators. Secondly, each project manager reviewed the translations delivered by its team before submitting the files and, in a subsequent stage, all translations were again cross-revised by other groups. Finally, after all groups had delivered their translations, an instance of the PKP software running on the university's servers was updated with the translated files. This allowed the students to get to know what a real testing process on localised software is like. Students were therefore asked to crawl the software, capture any kind of errors they could come across (linguistic, graphical, functional, etc.) using screen-shot software, and document the errors' nature through a classification template. This template was used to correct some linguistic issues and sent as feedback to PKP contributors.

6 Concluding Remarks

In this paper we have presented how openness is becoming more and more a key concept on translation following our translation project at the Tradumàtica Masters as a case in point. As mentioned earlier, we believe that FOSS software gives translation trainers an opportunity to teach how real localisation is carried out, overcoming ethical concerns and easing open access to knowledge to a greater community, thus becoming an asset transferred to society.

As this is a long-term, running project, year after year changes and modifications are included in its design. Some of the future working lines might include translation and translated software. Firstly, as for translation software, we attempt to include the latest technologies – with an eye on free software – to the workflow. In this sense, some technologies like Customised Machine Translation engines or proxy-based localisation might be researched; as of the academic year 2014-2015, the XLIFF standard has been included in the project design, following our belief that, as Jiménez-Crespo (2013: 176) puts it, "basic knowledge of exchange standards" is part of the technological subcompetence. Secondly, as for the translated products, other branches of the PKP software or even other products might be explored at some point, since it can be expected that, being somewhat similar and sharing files to some extent, a number of the chains will already be translated and stored in our translation memories.

References

Diaz Fouces, O. (2011) ¿Merece la pena introducir el software libre en la formación de traductores profesionales? Presented at *Language and Translation Teaching in Face-to-Face and Distance Learning*, Universitat de Vic. Vic, 8 April 2011. Available at http://www.academia.edu/3487697/_Merece_la_pena_introducir_el_software_libre_en_la_formacion_de_traductores_profesionales [Accessed 10 March, 2015].

Kancewicz-Hoffman, N. (2008) Increasing visibility for a multifaceted Humanities research in Europe – the ERIH approach. Presented in the Conference *Relevance and Impact of the Humanities*, Universität Wien. Vienna, 16 December 2008. Available at http://www.qs.univie.ac.at/fileadmin/user_upload/qualitaetssicherung/Veranstaltungen/Humanities/Praesentationen/nicht_bestaetigt/Kancewicz_Hoffmann_ERIH_impact_Vienna_Dec08.pps [Accessed 10 March, 2015].

Kelleher, M. and Hoogland, E. (2008) *Changing Publications Cultures in the Humanities*. Young Researchers Forum. ESF Humanitites Spring 2011. European Science Foundation – Humanities Unit. 9-11 June 2008, Maynooth, Ireland. Available at http://www.esf.org/fileadmin/Public_documents/Publications/Changing_Publication_Cultures_Humanities.pdf [Accessed 10 March, 2015].

Jiménez-Crespo, Miguel A. (2013) *Translation and web localization*. London, New York: Routledge.

Martín-Mor, A., Piqué Huerta, R. and Sánchez-Gijón, P. (forthcoming) *Tradumàtica. Tecnologies de la Traducció*. Vic: Eumo.

Munday, J. (2008) *Introducing Translation Studies: Theories and Applications*. London, New York: Routledge.

Open Journal Systems (n.d.) *Public Knowledge Project*. Available at https://pkp.sfu.ca/ojs/ [Accessed 10 March, 2015].

Piqué Huerta, R. and Sánchez-Gijón, P. (2013) Troubleshooting: el proyecto de formación de localizadores. *VI Congreso Internacional de la AIETI: Traducimos desde el sur*. Las Palmas de Gran Canaria, 23-25 January 2013.

Softcatalà (2010) *Guia d'estil*. Available at http://www.softcatala.org/wiki/Guia_d%27estil/Guia_2010 [Accessed 10 March, 2015].

To be or not to be a Scientist 2.0?
Open Access in Translatology
A German Case Study

Marco Agnetta
Saarland University, Germany

1 Introduction

Connoisseurs of linguistic mechanisms will not like the expression "scientist 2.0" which is employed in the title of the present study. This metaphor suggests that such a scientist would be an updated and ameliorated version of a sort of antiquated scientist 1.0. Although chosen as a provocative springboard, however, the question ("to be or not to be a scientist 2.0?") gets to the heart of a set of problems that arise out of presently changing scientific practices. Thus, why not begin with such a polemical wording in the title?

In recent years, a new conception of scientific activity for the 21st century has been put forward under the heading of "Open Science". This movement follows the recommendations formulated by the *Budapest Open Access Initiative* (BOAI 2001) and the *Berlin Declaration on Open Access to Knowledge in the Sciences and Humanities* (Berlin Declaration 2003) urging academic actors to ensure unrestricted access to knowledge, at least to that produced by themselves. In this context "Science 2.0" would mean the possibility (or utopian ideal?) of openly accessing any kind of knowledge resources produced or elaborated by researchers. "To be or not to be a Scientist 2.0?" is, therefore, a question that is becoming increasingly urgent in many disciplines, including also Contrastive Linguistics and Translatology. Paradoxically, this is occurring even though the indispensable adjustments specific to these disciplines that would follow from a positive response to the question have so far been neither defined nor applied. Nevertheless Open Access (OA) is flatly considered a revolutionary research practice (cf. Aschenbrenner et al. 2007: 21).

The present study does not try, nor is it able, to provide comprehensive solutions for these points of OA publishing which, more than a dozen years after the formulation of the above mentioned manifestos, are still denounced in our discipline. Within the framework of this study we will focus on the point of view of the academic actors on this new research and publication paradigm and we will investigate whether and to what extent realizations of OA endeavors can be found in contemporary German translatology. We will, therefore, explicitly refer to the activity of translation scholars and not to that

of translators or interpreters, where OA has also been identified as a significant desideratum (cf. further literature in this volume).

2 Openness in Translatological Research

In the Internet age open access is a frequently and vehemently voiced request which heavily affects conventional production and marketing conditions; this equally applies to public funded research. This is, inter alia, proved by the constantly increasing number of institutions that commit to the OA principle (cf. the *Registry of Open Access Repository Mandates and Policies*, ROARMAP). Despite its status as a ubiquitous expression in public and research discourse, openness must always be exactly defined. In general, one can speak of open access where barriers between customers or users and their product of interest do not exist: openness is equal to freedom from barriers. The *Open Knowledge Foundation* (OKFN) gives a more concrete definition of openness with regard to knowledge and mentions the following three "key features of openness" (cf. OKFN n.d.):

- "**Availability and access**: the data must be available as a whole and at no more than a reasonable reproduction cost, preferably by downloading over the internet. The data must also be available in a convenient and modifiable form.
- **Reuse and redistribution**: the data must be provided under terms that permit reuse and redistribution including intermixing with other datasets. The data must be machine-readable.
- **Universal participation**: everyone must be able to use, reuse and re-distribute – there should be no discrimination against fields of endeavour or against persons or groups. For example, 'non-commercial' restrictions that would prevent 'commercial' use, or re-strictions of use for certain purposes (e.g. only in education), are not allowed" (ibid.).

These points can be summarised to the following succinct definition formula propagated by the OKFN: "Open data and content can be freely used, modified, and shared by anyone for any purpose" (Opendefinition n.d.). This definition, as well as a more verbose version of it, are presently available in 38 languages (cf. ibid.). To comply with this definition of openness, persons and institutions who make available any kind of information and knowledge should, therefore, remove the following types of barriers:

1) *Access barriers*: These arise when gaining full or partial access to goods and services, whatever their nature, is inhibited by any spatial

and temporal conditions. We speak about technical barrier if we refer to the reduced accessibility to a certain medium.

2) *Pay/price barriers*: These arise when the access to and the use of goods and services is associated with monetary or any other considerations. Subscriptions, licensing fees, pay-per-view fees are current price barriers in scholarly publishing.

3) *Permission barriers*: These arise when the access to and the use of goods and services is fully or partially inhibited by legal regulations which specify manners and purposes of their utilization.

Herb (cf. 2015: 10-15) has already pointed out that openness is differently defined within the scientific community, where OA still means the removal of pay barriers for research output only. The accessibility to other information items like primary research data and software implemented for purposes of research is hardly ever granted. Scholars thus essentially content themselves with the definition of openness proposed by the BOAI (2001) that, according to Herb (2012: 11; 2015: 23), satisfies "minimum requirements" only. That is why he recommends the consistent terminological and conceptual distinction between "free" or "gratis" and "open" information items (cf. Herb 2015: 31-34). As we refer to the accessibility of scientific results only and not to their unrestricted re-use, we will subsequently work with the conventional proposition formulated as follows by Björk et al. (2013: 237): "literature that is merely free without granting liberal re-usage rights is still considered OA". Peter Suber, one of the best-known advocates of OA publishing, calls this kind of texts "royalty-free literature" and refers to them as very "low-hanging fruit of OA" (cf. Suber n.d.).

3 Open Access and the Research Cycle

At this point it is necessary to return to a chart of the research cycle as previously outlined by Agnetta (2015: 14-28). This description of research workflow will be completed with an analysis of the contemporary research and publication landscape in translatology. For this purpose a corpus of 115 explicit translation-related scientific journals (translating, interpreting or both) from all around the world and dating from 1995 until now has been compiled in order to examine whether and to what extent they conform to the OA principle (see Annex 1).

Academic activity of (comparative) philologists can be described as three successive and repeating phases: A. Research in a narrower sense, B. publication and C. the subsequent use of the generated or worked up knowledge. There is no categorical rejection of the OA principle in

contemporary humanities, as Agnetta has shown (cf. 2015: 13-14, 23). For scholars in the humanities already do make full use of all the benefits which go along with OA in the research paradigm (A.) (listed for instance in Fröhlich 1998: 545). Below we will follow up the extent to which the OA maxim is accepted in all of the above mentioned phases.
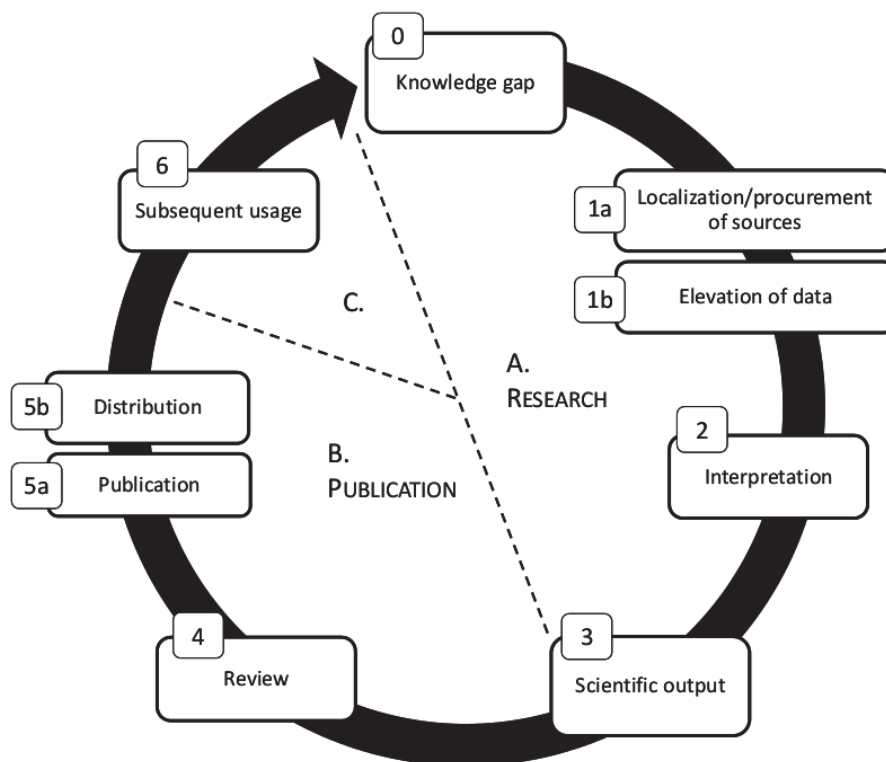
Figure 1: Research and publication workflow (Source: Agnetta 2015: 15).

(0) The research and publication workflow may be further divided into six single stations. It finds its starting point in the identification of a *knowledge gap* by one or more scholars while they are working with existent knowledge sources (be it printed or web media). It may be claimed that the more information is available without restrictions the more efficiently further knowledge gaps can be detected.

(1) With the aim of filling this knowledge gap, the philologist initiates his research including the localization and procurement of the sources (1a) and the acquisition of primary data (1b).

(1a) Localization and procurement of the sources: Online biblio-graphies, databases, and abstract services provide scholars with instruments which are presently indispensable for the localization of existing relevant literature and data. Those which can be fully or partially accessed in the Web can be located by means of certain Web services like *Google Scholar* or the *Bielefeld Academic Search Engine* (BASE). At best, these can be downloaded and printed as needed. Adema and Ferwerda (2009) debate whether OA makes sense for the publication of monographs which still dominate the humanities and social sciences ant they conclude that OA could "be a good alternative" (2009: 179) to conventional print publishing if determinate factors are taken into consideration. For the historical branches of translatology it is also one of the major goals that sources, at least those which are not protected by copyright, are available in digital scans or copies.

(1b) Elevation and procurement of the primary data: The success of many of the empirically working branches of Translation Studies depend on the availability of possibly already annotated corpora. Since their compilation is generally extremely time consuming and labor intensive, listings of searchable and possibly even workable corpora which include information about their free/open availability are of ever-increasing importance. This is one of the tasks of those centers of the Clarin-D consortium (Clarin-D n.d.) focusing mainly on (applied or comparative) linguistics as does for instance the *Hamburg Center for language corpora* (HZSK n.d.). Overviews over translatologically exploitable corpora are given for example in Possamai (2009) and Pontrandolfo (2012). In a research field with such an interdisciplinary orientation it is furthermore not negligible to which extent research results and data of neighboring disciplines are made available to Translatology.

(2) Interpretation: When primary and secondary sources have been procured they require quantitative and qualitative analysis. Here again institutions like Clarin-D provide corpus-based Translatology with infrastructures, tools and annotation criteria. According to the guidelines of the undermentioned CC-licensing, annotation is not included among those "derivates" that can be prohibited by the CC-ND-license (cf. Herb 2015: 20-21).

(3) Scientific output: On the basis of the sources' interpretation researchers put down in writing their results. In Translatology, monographs, contributions to collected volumes (in the form of conference papers and jubilee

publications), and to an increasing extent also journal articles are customary. In the humanities, where individual authorship remains the dominant mode of publishing, it is not usual to publish unfinished texts. Proofreading, exchange of views and quality control take place before formal publication. The dissemination of preprints is rarely found in these disciplines.

(4) Review: Journal articles and contributions to collected volumes generally pass through a multi-step reviewing procedure, in the course of which expert judgements are asked by the responsible editors. In the case of monographs, it is the post-publication recension that functions as equivalent "controlling instance" (Schütte [2]2009: 3). In the rest of the cases, pre-publication reviews ought to assure quality of the final and publishable manuscript. But it is precisely these reviewing procedures that are always accused of offering great manipulative potential because of the lack of transparency.

Herb (2010: 6ff.; 2012: 21-28; 2015: 169-195) discusses how far reviewing procedures should be made transparent for the whole scientific community by explaining new concepts of *collaborative* and *open reviewing*. Open reviews that name reviewer and reviewed scholar carry the risk of public humiliation of the latter since possible rejections would not only be visible, but also countable and finally evaluable. In the meantime, there are voices advocating at least a numerical publication and evaluation of generated reviews which are still not appreciated in common academic praxis, neither financially nor in terms of reputation. One initial approach to this purpose is presented by the website *Publons.com* (n.d.) that offers reviewers a platform to record their peer review contributions without breaking reviewer anonymity.

(5) Publication and distribution: After these multi-step quality assurance procedures the reviewed manuscript is sent to the publisher that has been commissioned for the formal publication (5a) and the distribution of printed or digital copies (5b).

(5a) Publication: The publishing landscape in translatology has significantly changed in the past two decades. Monographs (possibly in the form of doctoral or postdoctoral theses) and collected books find equal publication formats in the numerous OA journals. The online *Directory of Open Access Journals* (DOAJ) that compiles – albeit with some time lag – peer-reviewed OA journals from all over the world lists only two OA journals under the rubric "Translating and Interpreting" (as of August 2015). One more accurate search on the websites of the German *electronic journals database* (EZB n.d.) and Hispanic database *dialnet* (n.d.) provides a more comprehensive picture of existing translatological journals and their accessibility on the web:

Type \ Year	Total (in %)	founded before 2000 (journals before 2000)	founded between 2000 and 2014
Total	115 (=100%)	(47)	68
OA	78 (≈68%)	(22)	56
OA with restrictions	12 (≈2%)	10	2
non-OA	25 (≈30%)	15	10

Table 1: Journals in translatology.

This search yields a total number of 115 translatological journals published during the period between 1995 and 2014. Often it is no longer possible to reconstruct from which year certain print journals extended their offer by digitizing previous issues or by switching completely to OA publishing. Dates in brackets therefore do not necessarily refer to the publication type of a journal when it was established but rather to whether issues of those years are freely accessible from today's point of view. OA journals "with restrictions" are those restricting immediate open accessibility by any kind of non-disclosure notice or blocking period. All data given represents a snapshot dating August 2015.

Since 2000 not less than 56 translatological OA journals have been founded. And it should also be borne in mind that journals of related disciplines which could not be taken into account here provide a publishing platform for translation scholars as well. Foundations of journals which are not purely OA decrease more or less significantly after 2000. So it can be observed that more than two thirds of all existing translatological journals follow the OA maxim in 2015.

The question remains open whether authors are allowed to retroactively archive their printed articles in OA repositories (green road of OA publishing). According to information from the SHERPA/RoMEO database most of the publishers of non-OA journals only allow self-archiving or publishing of preprints or not copy-edited article versions which thus cannot be cited precisely. For journals which do not exist in this database (cf. column "not specified") it can be assumed that self-archiving is not welcomed either.

Archiving Type	Total	green publishing (not publisher's version)	yellow publishing (only pre-prints)	not specified (no self-archiving)
OA with restrictions	12	5	0	7
non-OA	25	7	1	17

Table 2: Self-publishing/archiving of articles in translatological journals

In the meantime many research institutes and research funders comply with the OA maxim and predicate financing on the condition that project-related publications should be made accessible in OA (cf. Herb 2015: 54-58). Detailed listings of such institutes and funders that have committed themselves to OA and which are mostly at the same time signatories of the above mentioned manifestos (BOAI, Berliner Erklärung) is provided by the SHERPA/JULIET database. According to this website, OA is – in Germany – explicitly encouraged or demanded in the publication guidelines of the *German Research Foundation* (DFG n.d.), the *Fraunhofer-Gesellschaft* and the *Helmholtz Association of German Research Centres*. These mandate the OA publication of research output (in the form of peer-reviewed original articles) and, in certain cases, even of primary research data (at the DFG). Free accessibility in appropriate repositories or the institute's own e-libraries (e.g. *Fraunhofer e-Prints*) is to be ensured as soon as possible, if need be when an imposed embargo period of six to eighteen months expires. However, important German research institutes and funders, even those which have decisively promoted the OA movement in Germany, have been omitted in this database, as has the Max Planck Society (n.d.) and the *Leibniz Association* (n.d.).

(5b) Distribution: More and more frequently researchers complain that most publishers merely seek to make a profit from the researchers' many years of work. Presently seen as mere money machines, publishers seem to have moved away from their original function of ensuring access to high quality research. Occasionally one can find extreme cases in which the content of volumes put on the market does not play any role if title and author (team) promise high turnovers. Assertions such as that quality is to be assured by publishers do not reflect reality – at least, not in the humanities. In the majority of cases, it is the authors themselves or the unpaid reviewers who bear responsibility for ensuring the absence of errors of content and form and who worry about editing and layout. Nevertheless, there is no need to condemn all existing publishers, since

several of them are beginning to extend their offerings by also establishing OA series.

However, it is important to mention that, especially in the case of OA journals conceived as such from the outset (golden road of OA publishing), costs are shifted from the recipient's to the producer's side, which means that author and potential funders now pay for publishing. The problem of social disadvantage frequently referred to in open OA discourse is now reproduced on the author's side: Whoever has the most money, publishes most. Alternative funding possibilities are described in Herb (2015: 60-82).

(6) Subsequent usage: Many entities are interested in the continued use of published research results, whether for again scientific, economic or simply individual information needs. It is undoubtedly a great achievement for OA movement that authors are able to retain the rights to the produced output as their intellectual property and to determine by themselves its further utilization. In recent times, *Creative Commons Licenses* (n.d.), which guarantee the naming of the author who has produced or elaborated the available contents (CC-BY), have become widespread in specifying the legal framework of subsequent usage of research results on the Web. In conventional publication workflows researchers were required to renounce their rights, ceding them to the publishing house they had chosen. Only a few publishers cede to the authors the right to archive their scientific output – after an embargo period of twelve to eighteen months from print publication – in appropriate repositories. In any case authors have to claim the contractual termination of such permission.

However, one fact in OA publishing is still considered a serious problem and that is the long-term availability of digital objects, which is regarded as insufficient among many web users, researchers included. The above mentioned time barrier is cited here. In any case, there are several approaches for its removal. One of them consists in the open source system LOCKSS (*Lots Of Copies Keep Stuff Safe*, n.d.) which ensures the long-term preservation of digital contents by means of their sevenfold storage in locally separated and hard drives (LOCKSS boxes) distributed all over the world. This prevents information loss in the case one or more hard drives fail. Questions concerning one binding standard electronic format for scientific results, as requested by the *Berlin Declaration* (2003), still remain unresolved.

4 Open Access and Academic Practice

Up to here our statements have been contingent on one condition whose fulfillment cannot be assumed flatly among scientists: The researcher does

support OA! Some barriers to research results are involuntarily or not least voluntarily created by scholars to protect themselves from present-day hostile academic mechanisms.

4.1   Open Access in University Education

An unsatisfactory system at universities for raising the level of awareness concerning publication possibilities and alternatives can be considered one of the involuntarily existing barriers to open accessibility. It may thus be argued that there is a genuine need for awareness campaigns.

We may assume that future translatologists first come into contact with the discipline during their time at university and that one of their first publishing experiences is the publication of a university thesis. A study attempting to explore how far the opportunity for OA publishing is available to German translatologists from the outset of their career should therefore commence with higher education institutes.

An in-depth analysis of the repository landscape in the German-speaking area is provided by the "2014 Census of Open Access Repositories in Germany, Austria and Switzerland" (cf. Vierkant/Kindling 2014). This statistical survey reveals that 42.01% of all universities (artistic higher education institutes included) and 9.38% of all technical colleges on German territory do operate OA repositories. In this context, the Göttingen State and University Library (*SUB Göttingen*) deserves particular mention due to the fact that this institution has committed itself to the setting up and maintenance of digital research environments and research infrastructures for data and services.

In the following it has to be established whether (young) German translatologists have the opportunity to publish their theses (BA, MA, doctoral and postdoctoral theses) in such repositories. Therefore, all state universities have to be listed, at least in terms of numbers, in which studies in translatology can be taken up. In a relevant German manual (*Handbuch der Universitäten und Fachhochschulen*, HUF [22]2012), seven universities and technical colleges are listed under the search items "translatology" and "interpretation/translation". This listing has been updated and complemented through our own investigation (see Annex 2). Half of the total of fourteen identified higher education institutes offer the opportunity to pursue a doctorate or habilitation. With the aid of the online *Registry of Open Access Repositories* (ROAR, n.d.) and our own web search it was possible to verify whether the respective education institution operates a publication server and/or OA publisher of its own. 13 of the 14 higher education institutions offer the possibility of OA publication of at least doctoral theses; the only exception is one technical college. If we refer to the above mentioned Census (2014), this result corresponds to the normal

case. It therefore can be proved that young translatologists of nearly all higher education institutions in Germany have the opportunity of OA publication.

But a broader awareness campaign still remains desirable. OA publication as an alternative to conventional book publishing could be explicitly integrated in examination, doctorate, and habilitation regulations in the humanities. In this regard, initiatives of three German universities play a pioneering role: These are on the one hand the cooperation program *MAP – Modern Academic Publishing* (n.d.) between the universities of Cologne and Munich and on the other the OA publisher of Saarland University *universaar* (n.d.).

Congress organizers could also be strongly encouraged to support OA publishing of the collected conference papers. One example of this may be the EU-financed translatological conference series on "Multidimensional Translation – MuTra" held in Saarbrücken (2005), Copenhagen (2006), and Vienna (2007), whose proceedings are entirely available on the Web. All of the OA publishing researchers have furthermore the choice to let their works (to which they retain all rights) be printed and marked by external and independent print-on-demand service providers like *Monsenstein und Vannerdat* or *Epubli*. Such hybrid publication models will surely become increasingly attractive in the future.

4.2   Academic Practice, Scientometrics and Open Access

Answers to the question whether OA and Open Science are largely accepted within the scientific community must necessarily take into account the structures and functioning of university career paths (cf. Agnetta 2015: 13). One could suppose that younger researchers support OA rather than established scholars since the former are often more technophilic and call into question the strict hierarchical academic structures. But this is not the case in times like these.

Anyone who imprudently publicizes Open Science as a common ideal will quickly be confronted with the utopian character of such a perspective. Even if Suber (2015) proves that "to advance knowledge does not conflict with the strong self-interest in career-building", it may be argued that OA to and altruistic provision of information seems to be undesired wherever research results promote academic or economic competitiveness. Non-disclosure notices specified by clients from economy and politics and the voluntary shortage or detention of research data by academic actors are no surprise within a context of competitive thinking and performance pressure. This concerns the humanities as much as the natural sciences. The massive budgetary cutbacks recently recorded across Germany are surely not welcome in this respect either.

Job offerings, involvement in projects, etc., depend more and more on questionable performance measurements that consider only publication activity and third-party fundraising disregarding other academic activities, teaching above all. Therefore, it is no surprise that research and publishing activity of scholars results partially from extrinsically motivated decisions, which means that they are not immediately related to the purpose of scientific progression (cf. Merton 1988: 621). That is why philosophers of science like Fröhlich call into question the intention of scientists to communicate optimally with their colleagues. He proves that retention, blockage, and retardation of information are current "effective strategies" even in the same research institution (cf. Fröhlich 1998: 536). If, on the other hand, proponents of OA accuse scientist of ignoring OA discourse within their own research, it may be replied that for many researchers this would mean a further distraction from the own research interest.

And thus emerges the quite paradoxical situation in which younger researchers have less interest in the open and free accessibility of their research results than established senior researchers. Thereby we want to address the importance of central institutions, whose task should be to provide, preserve and optimize functioning infrastructures for science in continuous consultation and cooperation with researchers.

Fröhlich (1998: 544ss.) paints a sobering picture: OA principle and web communication hold the potential to democratize science. But changing the problematic issues we have just touched on is not inevitably connected to changing the medium of publication. Existing problems will not suddenly be abolished if scholarship shifts to OA publishing. In truth, cases will continue to exist in which OA research infrastructure proves to be as vulnerable to abuse as conventional print models were (currently in Spain: cf. Sánchez Perona 2015 and Aréchaga n.d.). The OA system has also been successfully challenged by provocative researchers (cf. *scholarlyoa.com* n.d. and *SCIgen* n.d.). A gift economy based on reciprocity can be set up on the web as well as in non-web-based research environments by replacing mutual citation with interlinking for example (cf. Fröhlich 1998: 539-40).

It remains, thus, questionable whether in the future platforms will prevail which explicitly claim a return to research ethics and which offer scholars an environment in which they can do their research detached from extrinsic considerations, as the website www.sjscience.org holds out the prospect of.

4.3   Linguistic Diversity as Symptom of Research Diversity

There is general acknowledgement that all communication in the (natural) sciences should not be culture-specific, and the humanities also basically

endeavor to achieve intersubjectivity and intercomprehensibility. In view of the continuing internationalization of science there is one implicit request scholars feel themselves confronted with: It consists in the fact that they have to publish their works in English in the interests of increased visibility.

This may not be seen as problematic by OA supporters since a binding use of English as the *lingua franca* of science would mean the removal of an additional barrier to knowledge resources: that of the language. It need not be explained that English appears best-suited to take on the function of language of science by virtue of the number of (non-native) speakers. There are also linguistic peculiarities of English such as its practicability and simpler learnability that definitely suggest its use as common language in science (cf. Stackelberg 1988/2009: 5).

However, particularly in the philologies, in comparative linguistics, and translatology such demands cause a lot of contention. For many philologists equate research diversity with language diversity. It is in this spirit that Jürgen von Stackelberg, German Romance philologist and comparatist, defends the fact that scholars only meet the requirements of the own research subject when they draft their research results in their native language (cf. Stackelberg 1988/2009: 22). He views this trend towards making scientific research solely available in English as extrinsically motivated behavior on the part of researchers: "Humanists do, therefore, obey 'external' constraints. There are other than science immanent reasons when they publish in English" (ibid.: 10, translation: M.A.).

English is the most widely represented language in the submissions guidelines of the journals of our corpus (see Table 3). Other "major" languages are accepted in less than 50% of cases, but at the same time the percentage of pure OA journals is much higher in these languages than in English.

Total 115 Journals	Total (language)	% of Total	OA (in %)	not/partially OA (in %)
English	96	83%	65 (68%)	31 (32%)
French	47	41%	40 (85%)	7 (15%)
Spanish	45	39%	37 (82%)	8 (18%)
German	23	20%	19 (83%)	4 (17%)
Portuguese	20	17%	20 (100%)	0 (0%)
Italian	17	15%	15 (88%)	2 (12%)
Catalan	8	7%	8 (100%)	0 (0%)
Serbian	3	3%	3 (100%)	0 (0%)

Total 115 Journals	Total (language)	% of Total	OA (in %)	not/partially OA (in %)
Chinese	2	2%	1 (50%)	1 (50%)
Russian	2	2%	1 (50%)	1 (50%)
Dutch	1	< 1%	1 (100%)	0 (0%)
Galician	1	< 1%	1 (100%)	0 (0%)
Japanese	1	< 1%	0 (0%)	1 (100%)
Korean	1	< 1%	0 (0%)	1 (100%)
Norwegian	1	< 1%	1 (100%)	0 (0%)
Polish	1	< 1%	0 (0%)	1 (100%)
Romanian	1	< 1%	1 (100%)	0 (0%)
X: language not specified or 'further languages': 5 – 4% – 4 (80%) -– 1 (20%)				

Table 3: Languages in translatological journals.

Even though it is clear that what Stackelberg says results from a deep but individual conviction and one can find only few rational points in his argumentation, such statements bear witness to the great reservations many other philologists express with respect to anglicization of science language. Such voices are becoming loud in other countries, too, as is happening in Italy and France. In an issue of the French magazine *Circuit – Le magazine d'information des langagiers* (41/September 1993) that focuses on this topic (Title: *L'Europe au rythme de l'anglais*) Cormier/Humbley (1993: 2) worriedly observe that 80% of all scientific texts are already drafted in English (cf. also the satirical contribution "How did science come to speak only English" by Michael D. Gordin 2015). That communication and cooperation across borders is essential for research is in no case disputed by humanities scholars. But many of them agree that the binding use of English as the only *one* "langue véhiculaire" (Cormier/Humbley 1993: 2) is appropriate for texts of mere administrative character (reports and announcements for instance) or for the overwhelming majority of publications in the natural sciences but it is undesired in humanities and arts (cf. Stackelberg 1988/2009: 5, 11).

One might accuse Stackelberg of having a naive view of language when he suggests that institutions could impose the use of one common language on researchers. After all, language history proves impressively that normative language imposition is always shattered sooner or later. According to Stackelberg (1988/2009: 7) the intention to implement the use of a common language in science would, therefore, be an anachronism. And yet the

reservations formulated by the not primarily anglophone scientific community are not entirely unfounded.

In those disciplines in which quantifiable indicators are supposed to give information about research quality the use of English becomes, even if not explicitly stated, a necessary precondition for being noticed and cited outside the confined national borders. Besides third-party fundraising, citation remains the most important indicator for performance evaluation in research. The French anglicist Pierre Truchot (1993: 7) gets to the heart of the matter by formulating: "l'anglais ou l'anonymat" (English or anonymity). The demand for international comparability and the scientometrical analyses presently perform the function of a language standardizing institution.

So it is no wonder that journals of non-anglophone countries almost exclusively publish articles in English, as does the German OA journal *TC3 – Translation: Computation, Corpora, Cognition*. At least, one concession is made to the intrinsic multilingualism of translatology when "one paper per issue which is written in a language other than English" is accepted.

The preference for English submissions, abstracts and data mining is justified by the increased visibility of the scientific output. However, this is not the only reason. The translatological OA journal *Hermēneus* (n.d.) that accepts at least five languages apologizes to the submitters of differing linguistic skills that "experts with the proper linguistic competence and knowledge in pertinent fields in languages other than those mentioned are not often available to evaluate articles". In a young discipline such as translatology which has numerically far fewer scholars than other sciences, availability of experts that allow quality assurance of contributions in the minor language simply cannot be guaranteed.

We thus agree with Stackelberg (1988/2009: 4, 22) when he notes that the true removal of language barriers can only be initiated by means of translations. Also the OA journal from our corpus, *452ºF: The Journal of Literary Theory and Comparative Literature* agrees with this view by committing itself to multilingualism, to "[s]atisfy the need of a multilingual world: relying on the intrinsic cultural value of linguistic diversity, together with the need to reach as many readers as possible, several linguistic barriers will be avoided" (452ºF n.d.).

Good translation of reliable scientific literature might in future meet with the same academic appreciation as recensions and the preparation of didactic literature on the subject currently do. Anglophone research has already recognized this fact, as one can see from the language policies of the OA journal *Metamorphoses*: A *Journal of Literary Translation* that take "as its mission the publication of quality English language translation of the most

interesting articles […] presently available only in their source language"
(n.d.). The Hispanic journal *MonTI – Monografias de Traduccion e
Interpretacion* accepts translations to all minor languages in the online edition
and tries to provide English versions of all submitted articles.

5 Conclusions

Research in the humanities and especially in translatology is still far from
being part of an "Open Research Web" which is portrayed as a worthwhile
goal by Shadbolt et al. (2006). This is only partially due to the not fully
developed infrastructures which could ensure open access to all information
items that accrue in the course of the research and publication workflow. For
the way has definitely been already marked out. In fact, slow development in
this direction results from manifold and partially competing economic,
scientific-political and individual interests pursued by authors, users, research
institutions, publishers and more.

The presented discipline-specific analysis demonstrates that translatology
is no straggler in the matter of open accessibility and that it has already
internalized many issues of the OA movement. The sharp increase of
translatological OA journals, the availability of linguistic primary data and
corpora on the Web as well as the possibility of OA publishing at nearly all
tertiary education institutions which offer courses of translation studies testify
to a drive for innovation in our discipline. Here hybrid models that equally
provide for printed and online versions of contents legitimately predominate in
the publication landscape of translatology.

6 References

452ºF (n.d.) The Journal of Literary Theory and Comparative Literature Available at:
      http://452f.com/index.php/en/entidad-editora12 [Accessed August 2015].

Adema, J. and Ferwerda, E. (2009) Open Access for Monographs. LOGOS: The Journal of
      the World Book Community 20/1, 176-183. DOI: 10.1163/
      095796509X12777334632708.

Agnetta, M. (2015) Technik, die begeistert?! Zur Open-Access-Debatte in der heutigen
      Sprach- und Translationswissenschaft. In C. Polzin-Haumann and A. Gil (eds.)
      Angewandte Romanistische Linguistik. Kommunikations- und Diskursformen im 21.
      Jahrhundert. St. Ingbert: Röhrig Universitätsverlag, 11-28.

Aréchaga, J. (n.d.) Open Access, un arma de doble filo para las revistas científicas.
      Available at: http://www.sebbm.com/revista/articulo.asp?id=10007&catgrupo=265&
      tipocom=24 [Accessed August 2015].

Aschenbrenner, A., Blanke, T., Dunn, S., Kerzel, M., Rapp, A. and Zielinski, A. (2007) Von
      e-Science zu e-Humanities – Digital vernetzte Wissenschaft als neuer Arbeits- und

Kreativbereich für Kunst und Kultur. Bibliothek. Forschung und Praxis 31/1, 11-21. DOI: 10.1515/BFUP.2007.11.

BASE (n.d.) Available at: http://de.base-search.net [Accessed August 2015].

Berlin Declaration 2003 Available at: www.openaccess.mpg.de/Berliner-Erklaerung [Accessed August 2015].

Björk, B., Laakso, M., Welling, P. and Paetau, P. (2014) Anatomy of green open access. Journal of the Association for Information Science and Technology 65/2, 237-250. DOI: 10.1002/asi.22963.

BOAI 2001 Available at: www.budapestopenaccessinitiative.org [Accessed August 2015].

Clarin-D. (n.d.) Available at: http://www.clarin-d.de/en/home [Accessed August 2015].

Cormier, M. and Humbley, J. (1993) L'anglais en Europe: la croisée des Chemins. Circuit. Magazine d'information sur la langue et la communication 41, 2. Available at: http://www.circuitmagazine.org/images/stories/documents/archives/CI_41_93.pdf.

Creative Commons Licenses (n.d.) Available at: http://creativecommons.org) [Accessed August 2015].

DFG (n.d.) Available at: http://www.dfg.de/en/research_funding/programmes/infrastructure/lis/funding_opportunities/open_access_publishing/index.html [Accessed August 2015].

Dialnet (n.d.) Available at: http://dialnet.unirioja.es [Accessed August 2015].

DOAJ (n.d.) Available at: https://doaj.org/ [Accessed August 2015].

EZB (n.d.) Available at: http://rzblx1.uni-regensburg.de/ezeit [Accessed August 2015].

Fröhlich, G. (1998) Optimale Informationsvorenthaltung als Strategem wissenschaftlicher Kommunikation. In Zimmermann, H. and Schramm, V. (eds.) Knowledge Management und Kommunikationssys-teme, Workflow Management, Multimedia, Knowledge Transfer. Proceedings of the 6. Internationalen Symposium on Information Science (ISI 1998), Prague, 3.-7. November 1998. Konstanz: UVK Verlagsgesellschaft mbH, 535-549.

Gordin, M. (2015) Science once communicated in a polyglot of tongues, but now English rules alone. How did this happen – and at what cost? Available at: http://aeon.co/magazine/science/how-did-science-come-to-speak-only-english/ [Accessed August 2015].

Herb, U. (2012) Offenheit und wissenschaftliche Werke: Open Access, Open Review, Open Metrics, Open Science & Open Knowledge. In Herb, U. (ed.) Open Initiatives: Offenheit in der digitalen Welt und Wissenschaft. Saarbrücken: universaar, 11-44. Available at: http://universaar.uni-saarland.de/monographien/volltexte/2012/87/ [Accessed August 2015].

Herb, U. (2015) Open Science in der Soziologie: Eine interdisziplinäre Bestandsaufnahme zur offenen Wissenschaft und eine Untersuchung ihrer Verbreitung in der Soziologie. Glückstadt: Verlag Werner Hülsbusch. DOI: 10.5281/zenodo.31234.

Hermēneus (n.d.) Available at: http://www5.uva.es/hermeneus/?p=254&lang=en) [Accessed August 2015].

HUF ([22]2012) Handbuch der Universitäten und Fachhochschulen. Deutschland, Österreich, Schweiz. Berlin, New York: De Gruyter Saur.

HZSK (n.d.) Available at: https://corpora.uni-hamburg.de/drupal [Accessed August 2015].

Leibniz Assotiation (n.d.) Available at: http://www.leibniz-gemeinschaft.de/infrastrukturen/open-access [Accessed August 2015].

LOCKSS (n.d.) Available at: http://www.lockss.org [Accessed August 2015].

MAP (n.d.) Available at: http://www.humanities-map.net [Accessed August 2015].

Max Planck Society (n.d.) Available at: http://openaccess.mpg.de [Accessed August 2015].

Merton, R. K. (1988) The Matthew Effect in Science II – Cumulative Advantage and the Symbolism of Intellectual Property. ISIS, 79, 606-623.

Metamorphoses (n.d.) Available at: http://www.artintranslation.org/ [Accessed August 2015].

MonTI (n.d.) Available at: http://dti.ua.es/es/monti/monti.html [Accessed August 2015].

MuTra 2005-2007: Available at: http://www.euroconferences.info/proceedings/proceedings.php?proceedings=1) [Accessed August 2015].

OKFN (n.d.) Available at: www.okfn.org/opendata [Accessed August 2015].

Opendefinition (n.d.) Available at: www.opendefinition.org/od [Accessed August 2015].

Pontrandolfo, G. (2012) Legal Corpora: an overview. Rivista Internazionale di Tecnica della Traduzione 14, 121-136. Available at: http://hdl.handle.net/10077/9783 [Accessed August 2015].

Possamai, V. (2009) Catalogue of Free-Access Translation-Related Corpora. Revista Tradumàtica 7/2009. Available at: http://www.fti.uab.cat/tradumatica/revista/num7/articles/09/9art.htm

Publons.com (n.d.) Available at: https://publons.com [Accessed August 2015].

ROAR (n.d.) Available at: http://roar.eprints.org/ [Accessed August 2015].

ROARMAP (n.d.) Available at: http://roarmap.eprints.org [Accessed August 2015].

Sánchez Perona, J. (2015) La peligrosa deriva de las publicaciones en acceso abierto. Available at: http://cienciaconfuturo.com/2015/07/23/la-peligrosa-deriva-de-las-publicaciones-en-acceso-abierto/ [Accessed August 2015].

scholarlyoa.com (n.d.) Available at: http://scholarlyoa.com/2014/11/20/bogus-journal-accepts-profanity-laced-anti-spam-paper [Accessed August 2015].

Schütte, G. ([2]2009) Zählen, gewichten, lesen. Zur Bewertung von wissenschaftlichen Publikationsleistungen in Peer review-Prozessen. In Alexander von Humboldt-Stiftung (ed.) Publikationsverhalten in unterschiedlichen wissenschaftlichen Disziplinen. Beiträge zur Beurteilung von Forschungsleistungen, 212/2009, 3-4.

SCIgen (n.d.) Available at: http://pdos.csail.mit.edu/scigen [Accessed August 2015].

Shadbolt, N., Brody, T., Carr, L. and Harnad, S. (2006) The Open Research Web: A Preview of the Optimal and the Inevitable. In N. Jacobs (ed.) Open Access: Key Strategic, Technical and Economic Aspects. Oxford: Chandos Publishing, 13-26. Available at: http://eprints.ecs.soton.ac.uk/12453/ [Accessed August 2015].

SHERPA/JULIET (n.d.) Available at: www.sherpa.ac.uk/juliet [Accessed August 2015].

SHERPA/RoMEO (n.d.) Available at: http://www.sherpa.ac.uk/romeo [Accessed August 2015].

Stackelberg, J. (1988/2009) Künftig nur noch Englisch? Ein Plädoyer für den Gebrauch der Muttersprache in den Geisteswissenschaften. Bonn: Romanistischer Verlag Jakob Hillen.

Suber, P. (2015) Open Access Overview. Focusing on open access to peer-reviewed research articles and their preprints. Available at: http://legacy.earlham.edu/ ~peters/ fos/overview.htm [Accessed August 2015].

Truchot, C. (1993) La communication scientifique en Europe : l'anglais ou l'anonymat. Circuit. Magazine d'information sur la langue et la communication, 7-8. Available at: http://www.circuitmagazine.org/images/stories/documents/archives/CI_41_93.pdf [Accessed August 2015].

Universaar (n.d.) Available at: http://www.uni-saarland.de/campus/service-und-kultur/ medien-und-it-service/universaar.html [Accessed August 2015].

Vierkant, P. and Kindling, M. (2014) Welche Institutionen betreiben Open-Access-Repositorien in Deutschland? LIBREAS. Library Ideas. Available at: http://libreas.eu/ausgabe26/ 07vierkantkindling/ [Accessed August 2015].

Annex 1: OA Journals in Translatology

In the following we present our corpus of 115 explicit translation-related scientific journals (translating, interpreting or both) from all around the world and dating from 1995 until now. It has been compiled in order to examine whether and to what extent they conform to the OA principle.

1. 1611: Revista de Historia de la Traducción
2. 452ºF, The Journal of Literary Theory and Comparative Literature
3. Across Languages and Cultures
4. Alternative Francophone
5. Art in Translation
6. Asia Pacific Translation and Intercultural Studies
7. Babel
8. Babilónia: Revista Lusófona de Línguas, Culturas e Tradução
9. Between
10. Bulletin du CRATIL
11. Cadernos de Literatura em Tradução
12. Cadernos de Tradução
13. Circuit : Magazine d'Information sur la Langue et la Communication
14. Communication and Culture Online
15. Compilation and Translation Review
16. Computers and Translation
17. Confluências : Revista de Tradução Científica e Técnica
18. Critical Multilingualism Studies
19. Cultura e Tradução
20. Cultural Intertexts
21. Doletiana: Revista de Traducció, Literatura i Arts
22. Entreculturas
23. Estudios de Traducción
24. Eutomia : Journal of Literature and Linguistics
25. Forfatteren Oversetteren
26. Hermeneus: Revista de la Facultad de Traducción e Interpretación de Soria
27. Hieronymus complutensis. El mundo de la traducción
28. Hikma: Estudios de traducción
29. J-ELTS, International Journal of English Language and Translation Studies
30. In other words
31. Interculturalidad y traducción. Revista internacional
32. International Journal of Interpreter Education
33. Interpreting
34. In-Traduções. Revista do Programa de Pós-Graduação em Estudos da Tradução da UFSC
35. InTRAlinea : Online Translation Journal
36. JoSTrans: The Journal of specialised Translation
37. Journal of Applied Linguistics and Language Research
38. Journal of Interpretation Research
39. Journal of King Saud University - Languages and Translation
40. Journal of Translation
41. Koiné. Quaderni di ricerca e didattica sulla traduzione e l'interpretazione
42. La Linterna del Traductor
43. L'Antenne Express
44. Lebende Sprachen
45. L'Écran Traduit
46. Linguaculture
47. Linguística : Revista de Estudos Linguísticos da Universidade do Porto
48. Linguistica Antverpiensia. New series. Themes in Translation Studies
49. Livius.Revista de estudios de traducción
50. Machine Translation
51. Machine Translation Review
52. Meta: Journal des Traducteurs
53. Metamorphoses: A Journal of Literary Translation

54. Między Oryginałem a Przekładem
55. MonTi. Monografás de Traducción e Interpretación
56. Mutatis Mutandis. Revista Latinoamericana de Traducción
57. New Voices in Translation Studies
58. Norwich Papers
59. Língua – Revista Digital sobre Tradução
60. Onomázein : Revista de Lingüística, Filología y Traducción
61. Palimpsestes. Revue de Traduction
62. Panace@ [Panacea]: Boletín de Medicina y Traducción
63. Papers Lextra: Revista electrònica del Grup d'Estudis Dret i Traducció
64. Perspectives : Studies in Translatology
65. Philologia
66. Professional Communication and Translation Studies
67. Puentes: Hacia nuevas investigaciones en la mediación intercultural
68. Pusteblume. Journal of Translation
69. Quaderns: Revista de Traducció
70. Recherches et Travaux
71. Redit, Revista Electrónica de Didáctica de la Traducción y la Interpretación
72. Revista de Lingüística y Lenguas Aplicadas
73. Revista Tradumàtica : Traducció i Tecnologies de la Informació i la Comunicació
74. Rivista Internazionale di Tecnica della Traduzione
75. Saltana
76. Scientia Traductionis
77. Sendebar
78. Senez
79. Skopos : revista internacional de traducción e interpretación
80. Studii de gramatică contrastivă
81. T21N : Translation in Transition
82. Target
83. TC3 - Translation : Computation, Corpora, Cognition

84. TEXTconTEXT
85. The Bible Translator
86. The interpreter's Newsletter
87. The Journal of Interpretation
88. The Translator. Studies in Intercultural Communication
89. Ticontre: Teoria, Testo, Traduzione
90. Trabalhos em Lingüística Aplicada
91. Traces. A multilingual journal of cultural theory and translation
92. TradTerm
93. Tradução & Comunicação : Revista Brasileira de Tradutores
94. Tradução em Revista
95. Traducción & Comunicación
96. Traduction, Terminologie, Rédaction (TTR)
97. Traduire
98. Tradurre
99. Traduttologia
100. Trans : Revista de Traductología
101. Transfer. Revista Electrónica sobre Traducción e Interculturalidad
102. Trans-kom
103. Translation : A Transdisciplinary Journal
104. Translation and Interpreting
105. Translation and Interpreting Studies (TIS): The Journal of the American Translation and Interpreting Studies Association
106. Translation and Literature
107. Translation Journal: A Publication for Translators by Translators about Translators and Translation
108. Translation Review
109. Translation Spaces
110. Translation Studies
111. Translation Today
112. Translation Watch Quarterly: A Journal of Translation Standards Institute
113. Translationes
114. Two Lines – A Journal of Translation
115. Viceversa: Revista galega de traducción

Annex 2: OA in German State Universities

In the following, all state universities have been listed, at least in terms of numbers, in which studies in translatology can be taken up. In the German manual (*Handbuch der Universitäten und Fachhochschulen*, [22]2012), seven universities and technical colleges are listed under the search items "translatology" and "interpretation/translation".

1. **Fachhochschule Köln**: Fakultät für Informations- und Kommunikationswissenschaften; Institut für Translation und Mehrsprachige Kommunikation
   Fachübersetzen (Englisch, Französisch, Spanisch),
   Konferenzdolmetschen (Englisch, Französisch, Spanisch)
   Promotions- und Habilitationsmöglichkeit nicht gegeben
   OA: Cologne Open Science (http://opus.bsz-bw.de/fhk); Fachrepositorium (Informationswissenschaft): PubLIS Cologne (http://publiscologne.fh-koeln.de/home)

2. **Ruprecht-Karls-Universität Heidelberg**: Philosophische Fakultät; Institut für Übersetzen und Dolmetschen (IÜD)
   Übersetzungswissenschaft [B.A.] (Englisch, Französisch, Italienisch, Portugiesisch, Russisch Spanisch)
   Translation Studies for Information Technologies [B.A.] (Englisch)
   Übersetzungswissenschaft [M.A.] (Englisch, Französisch, Italienisch, Portugiesisch, Russisch Spanisch)
   Konferenzdolmetschen [M.A.] (Englisch, Französisch, Italienisch, Japanisch, Portugiesisch, Russisch, Spanisch)
   Promotions- und Habilitationsmöglichkeit gegeben
   OA: HeiDok – Heidelberger Dokumentenserver (http://archiv.ub.uni-heidelberg.de/volltextserver)

3. **Universität Hildesheim**: Fachbereich 3: Sprach- und Informationswissenschaften; Institut für Übersetzungswissenschaft und Fachkommunikation
   Internationale Kommunikation und Übersetzen [B.A.] (Englisch, Französisch, Spanisch)
   Medientext und Medienübersetzung [M.A.] (Englisch, Französisch, Spanisch)
   Promotions- und Habilitationsmöglichkeit gegeben
   OA: HilDok – Publikationsserver der Universität Hildesheim (http://hildok.bsz-bw.de/home)

4. **Universität Leipzig**: Philologische Fakultät; Institut für Angewandte Linguistik und Translatologie
   Translation [B.A.] (Englisch, Französisch, Russisch, Spanisch)
   Interkulturelle Kommunikation und Translation [B.A.] (Tschechisch-Deutsch)
   Translatologie [M.A.] (Englisch, Französisch, Russisch, Spanisch)
   Fachübersetzen [M.A.] (Arabisch, Deutsch)
   Konferenzdolmetschen [M.A.] (Arabisch, Englisch, Französisch, Russisch, Spanisch)
   Promotions- und Habilitationsmöglichkeit nicht gegeben
   OA: Qucosa – Publikationsserver der Universität Leipzig (http://ul.qucosa.de/startseite)

5. **Hochschule Magdeburg-Stendal** (Standort: Magdeburg): Fachbereich Kommunikation und Medien,
   I*nternationale Fachkommunikation und Übersetzen* [B.A.] (Deutsch, Englisch)
   Dolmetschen und Übersetzen für Gerichte und Behörden [Zertifikat, 2 Sem.] (je nach Nachfrage)
   Promotions- und Habilitationsmöglichkeit nicht gegeben
   OA: Digitale Hochschulbibliothek Sachsen-Anhalt [Universitätszusammenschluss] (https://www.hs-magdeburg.de/home.html)

6. **Hochschule für angewandte Sprachen München**:
   Internationale Technik- und Medienkommunikation [B.A.] (Englisch)
   Übersetzen [B.A.] (Chinesisch)
   Internationale Medienkommunikation [M.A.] (Englisch)
   Konferenzdolmetschen [M.A.] (Englisch)
   Promotions- und Habilitationsmöglichkeit nicht gegeben
   OA: nicht vorhanden, OA-Publikationsmöglichkeit nicht bekannt

7. **Universität des Saarlandes** (Standort: Saarbrücken): Philosophische Fakultät II; Fachrichtung 4.6, Angewandte Sprachwissenschaft sowie Übersetzen und Dolmetschen
   Vergleichende Sprach- und Literaturwissenschaft sowie Translation (VSLT) [B.A.] ((Englisch, Französisch, Italienisch, Spanisch): läuft aus
   Translationswissenschaft: Übersetzen [M.A:] (Deutsch (für Frankophone), Englisch, Französisch, Italienisch, Spanisch) läuft aus
   Translationswissenschaft: Konferenzdolmetschen [M.A:] (Deutsch (für Frankophone), Englisch, Französisch, Spanisch): läuft aus
   Promotions- und Habilitationsmöglichkeit gegeben
   OA: SciDok – Open-Access-Server (http://scidok.sulb.uni-saarland.de); OA-Verlag: universsar (http://www.uni-saarland.de/campus/service-und-kultur/medien-und-it-service/universaar.html)

This listing has been updated and complemented through our own investigation:

8. **Heinrich-Heine-Universität Düsseldorf**: Philosophische Fakultät; Institut für Romanistik
   Literaturübersetzen [M.A.] (Englisch, Französisch, Italienisch, Spanisch)
   Promotions- und Habilitationsmöglichkeit gegeben
   OA: Düsseldorfer Dokumenten- und Publikationsservice (http://docserv.uni-duesseldorf.de/)

9. **Fachhochschule Flensburg:**
   Internationale Fachkommunikation/Technikübersetzen [B.A.] (Deutsch, Englisch)
   Internationale Fachkommunikation/Technikübersetzen [M.A.] (Deutsch, Englisch)
   Promotions- und Habilitationsmöglichkeit nicht gegeben
   OA: e-Publikationsdienst: Zentrale Hochschulbibliothek Flensburg (http://www.zhb-flensburg.de/)

10. **Johannes-Gutenberg-Universität Mainz** (Standort: Germersheim): Fachbereich 06: Translations-, Sprach- und Kulturwissenschaft
    Sprache, Kultur, Translation [B.A.] (Arabisch, Deutsch, Englisch, Französisch, Italienisch, Neugriechisch, Niederländisch, Polnisch, Portugiesisch, Russisch, Spanisch, Türkisch)
    Translation [M.A.] (Arabisch, Chinesisch, Deutsch, Englisch, Französisch, Italienisch, Neugriechisch, Niederländisch, Polnisch, Portugiesisch, Russisch, Spanisch, Türkisch)
    Konferenzdolmetschen [M.A.] (Deutsch, Englisch, Französisch, Italienisch, Neugriechisch, Niederländisch, Polnisch, Portugiesisch, Russisch, Spanisch)

Promotions- und Habilitationsmöglichkeit gegeben
OA: ArchiMeD – Archiv Mainzer elektronischer Dokumente (http://archimed.uni-mainz.de/opusubm/archimed-home.html)

11. **Ludwig-Maximilian-Universität (LMU) München**: Fakultät für Sprach- und Literaturwissenschaften; Departament III: Anglistik und Amerikanistik
Literarisches Übersetzen [M.A.] (Englisch, Französisch, Spanisch, Italienisch)
Promotions- und Habilitationsmöglichkeit gegeben
OA: Elektronische Dissertationen der LMU München (http://edoc.ub.uni-muenchen.de/)

12. **Westfälische Wilhelms-Universität Münster**: Fachbereich 09: Philologien; Institut für Niederländische Philologie
Literarisches Übersetzen und Kulturtransfer (LÜK) [M.A.] (Niederländisch): läuft aus, stattdessen ab WS 2015/16: *Interdisziplinäre Niederlandistik* [M.A.]
Promotions- und Habilitationsmöglichkeit gegeben
OA: miami – Münstersche Informations- und Archivsystem multimedialer Inhalte (http://www.uni-muenster.de/Publizieren/dienstleistungen/repository/)

13. **Hochschule für angewandte Wissenschaften Würzburg-Schweinfurt** (Standort: Würzburg): Fachübersetzen und mehrsprachige Kommunikation
Fachübersetzen (Wirtschaft/Technik) [B.A.] (Englisch, Französisch, Spanisch)
Fachübersetzen und mehrsprachige Kommunikation [M.A.] (Deutsch, Englisch, Französisch, Spanisch)
Promotions- und Habilitationsmöglichkeit nicht gegeben
OA: FH-WS: Publikationsserver der Hochschule Würzburg-Schweinfurt (http://bibliothek.fhws.de/service/elektronisches_publizieren.html)

14. **Hochschule Zittau/Görlitz:** Fakultät Management und Kulturwissenschaften
Übersetzen [B.A.] (Englisch/Polnisch, Englisch/Tschechisch): läuft aus
Fachübersetzen Wirtschaft [M.A.] (Polnisch)
Promotions- und Habilitationsmöglichkeit nicht gegeben
OA: Qucosa – Der sächsische Dokumenten- und Publikationsserver (http://www.qucosa.de/startseite)

Digital Scholarship in Translation Studies: a Plea for Openness

Peter Sandrini
University of Innsbruck, Austria

Free and open source software defines openness with regard to the free availability of the source code and the binary program. Beyond free availability and gratuitousness, however, there is a more profound rationale behind the concept of openness, touching the question of social equality when referring to knowledge and education, as well as to the ownership of knowledge in general. The academic world, and researchers in particular, are at the core of this challenge which has intensified significantly with globalization tendencies and the digital revolution. Theoretically, principles and practice of academic work remain the same: researchers and scholars still strive for valid and trustworthy methods of inquiry. The environment in which studies are carried out, documented and published, though, has undergone deep changes. It provides new possibilities, linking the practice of scholarship with the possibilities of digital technology and new media. Digital scholarship has many dimensions and may be defined as "the use of digital evidence and method, digital authoring, digital publishing, digital curation and preservation, and digital use and reuse of scholarship" (Smith Rumsey 2013: 158).

The following paper concentrates on the concept of openness in the use of digital technology and digital media in academic research, and Translation Studies (TS) in particular, leaving aside the exploration of openness within two other important areas of digital scholarship: the use of digital technology in education and training, as well as the study and analysis of the digital medium itself.

To this end, we need to take a look at publication methods, access options to publications, as well as academic evaluation methods in TS, a research field where we have to deal with the peculiarity of different publication languages and a variety of competing research methods and theories.

It is evident that digital scholarship or the "scientist 2.0" as called by Agnetta (in this volume) cannot elude the problems and common trends of the new digital world, and openness seems to be one of them. Discussions about open source code, open knowledge, open content, open data, open education, etc. have lead the way to the question of openness in research, openness in publishing research results, or open access. This paper wraps up the situation in TS and makes a plea for openness since more openness could foster the discipline as a whole and move it towards a more unified and collaborative field of study.

1 Open Access Publishing

The statements in this paper are based on the following assumptions regarding research publications, even if they are taken for granted by a majority of researchers and aptly called 'truisms' by Blommaert (2014: 6):

- the main purpose of publishing is finding a readership;
- research doesn't make sense without publishing results;
- the less barriers between potential readers and research results the better reception and response from readers, colleagues and fellow researchers.

At the beginning of modern scholarship Aristotle stated in his Metaphysics 'All humankind by nature desires to know' and Wilinsky (2006) deduces: "As this desire is rightly identified, I believe, as part of our nature, it stands as a human right to know" (Willinsky 2006: 27). The right to know on the side of the public is complemented by the desire to communicate on the side of researchers, and publishing is the medium of choice for academia.

The field of publishing in TS is very heterogeneous and distributed over different countries and languages, a fact called by Gile (2015: 240) "the geographic, thematic and methodological fragmentation of TS". Different countries have developed diverse theoretical approaches, and very often language barriers prevent adoption and discussion of foreign theories. Nevertheless, the specific object of study as such represents "more of an interlingual, cross-cultural, interdisciplinary, and supranational subject of international interest" (Xiangdong 2015: 184). Referring to the first outline of the discipline published by James S. Holmes in 1972, Xiangdong then goes on: "The main research areas in Holmes's' map of TS, for example, theoretical studies, descriptive studies, translator training, translation aids, and translation criticism, are all topics of global interest" (Xiangdong 2015: 184). A common scientific basis as well as knowledge of seminal publications and the most important theoretical approaches, independently of the language in which they were originally written, all this constitutes a precondition for a sound subject field, and a prerequisite for an evolving discipline.

Furthermore, TS is not always recognized as an autonomous discipline, but rather subsumed under linguistics, comparative literature, philology or communication studies in general (Rovira-Esteva and Orero 2012, Gentzler 2014, Xiangdong 2015). These factors make TS a challenging discipline when it comes to research and evaluation: access to theoretical literature and publications is essential for the first, consideration of the peculiarities and idiosyncrasies of the subject field fundamentally important for the second.

What may keep researchers from accessing relevant literature is financial barriers, restrictions in place and time, as for example location and opening times in public libraries, availability of publications, etc. A first step in overcoming those barriers was the advent of the Web with new possibilities for independent publication of all kinds of texts, enabling at the same time Online Public Access Catalogs (OPACS) which made meta information on publications freely available. A second and more important step was the removal of legal and financial barriers by introducing new license models, such as, for example, the 'Copyleft' model of free software, or the 'Creative Commons' licenses, as well as open access publication models.

The definitions of Open Access (OA) are not always clear-cut or consistent: broad descriptions define OA as being found freely available online, others describe it as the "removal of barriers (including price barriers from accessing scholarly work" (Eysenbach 2006: 1). The founding papers and declarations of OA provide a more detailed description:

> "free availability on the public Internet, permitting any users to read, download, copy, distribute, print, search, or link to the full texts of these articles, crawl them for indexing, pass them as data to software, or use them for any other lawful purpose, without financial, legal, or technical barriers" (Budapest Open Access Initiative 2002).

For a work to be OA, the copyright holder must consent in advance to let users "copy, use, distribute, transmit and display the work publicly and to make and distribute derivative works, in any digital medium for any responsible purpose, subject to proper attribution of authorship" (Berlin Declaration 2003).

This is in open contrast to the copyright policies of commercial publishers who make researchers sign contracts which force them to hand over all rights to the publisher, in many cases even the right of re-use of published material, for example on a researcher's personal website. Such copyright agreements commonly impose severe restrictions on use while OA is the immediate, online, free availability of research output. The absence of legal barriers implies the existence of appropriate legal licenses. A suitable proposal has been developed by the Creative Commons (CC) framework shortly before the OA declarations, with the intention of creating a license model that enables people to "share your knowledge and creativity with the world" (creativecommons.org) in order to "maximize digital creativity, sharing, and innovation" (creativecommons.org). It offers six licenses based on a combination of the following rights modules: by (attribution), nc (non commercial), nd (no derivatives), sa (share alike), plus the public domain license CC0 (no copyright). As good practice in research already imposes, all six CC licenses require attribution of authorship; the nd restriction does not lend itself to research since

research heavily builds upon previous publications and it would be bad research if everybody should start anew from scratch.

It is precisely the fear of copyright violation, of lack of attribution, or the fear of unhindered stealing of ideas ('scooping') which keeps many scholars from embracing OA publication models although this is explicitly catered for by the different CC licenses. Yet, this reservation is expressed very often as an argument against OA, brought forward mainly by senior researchers who are not very familiar with new media. Being freely available, OA publications can be read and re-used by everyone, sometimes even copied illegally, but at the same time, any infringement on copyright can be easily identified through plagiarism checkers, even more so with OA online publications than with closed or restricted publications which are not always accessible to this kind of software checkers.

The main advantage of OA is the removal of obstacles between author and readers, opening up access for those who need it: scholars from small institutions and developing countries, patient advocates, patients themselves, and lay scholars. Basically, research and scholarly communication should be considered as a public good and publishing of research should be treated as such. Most research in translation is conducted by state-employed university staff paid for by the public. Thus, a certain moral obligation exists to make research outcome accessible to the public. Commercial publishers normally require authors to pay a publication fee which researchers usually take from institutional or public research funds, equally paid for by taxpayers, and then publishers charge the public, taxpayers again, money for the same publications in book form: thus, the public pays three times basically for the same research results.

John Willinsky, one of the world's leading advocates of OA, sees the free exchange of information as a matter of social justice, and estimates that already around 20-25 per cent of all peer-reviewed material currently published is now OA (Willinsky 2006).

Opening up readership means more readers who will read, process and absorb published ideas. An empirical study in physiology showed "full text downloads were 89% higher, PDF downloads 42% higher, and unique visitors 23% higher for open access articles than for subscription access articles" (Davis et al 2008), a result subsequently corroborated by another study involving 36 participating journals in the sciences, social sciences, and humanities, reporting that OA articles "received significantly more downloads and reached a broader audience within the first year, yet were cited no more frequently, nor earlier, than subscription-access control articles within 3 years" (Davis 2011: 2129), a finding confirmed elsewhere as well: "OA articles are

cited earlier and are, on average, cited more often than non-OA articles" (Eysenbach 2006: 696).

A larger readership results in increased uptake of research results and ideas, leading to a higher citation rate, indicating "that authors are finding them more easily, reading them more often, and therefore citing them disproportionately in their own work" (Antelman 2004: 377). The observation that OA articles receive more citations than subscription-based articles is known as the OA citation advantage (OACA): "it is clear that the advantage exists and occurs regularly across a range of subject areas" (Norris et al 2008: 1970). Eysenbach (2006) proposes a study with similar results in favor of OA publications for the subject field of biology, stating that "OA articles compared to non-OA articles remained twice as likely to be cited […] in the first 4-10 mo after publication […], with the odds ratio increasing to 2.9 […] 10-16 mo after publication" (Eysenbach 2006: 1). Another study (Antelman 2004) investigates

> "articles in four disciplines at varying stages of adoption of open access – philo-sophy, political science, electrical and electronic engineering and mathematics – to see whether they have a greater impact as measured by citations in the ISI Web of Science database when their authors make them freely available on the Internet. The finding is that, across all four disciplines, freely available articles do have a greater research impact" (Antelman 2004: abstract).

The website SPARC Europe lists 46 studies that found a citation advantage, 17 studies that found no citation advantage, and 7 studies "that were inconclusive, found non-significant data or measured other things than citation advantage for articles" (http://sparceurope.org/oaca/).

Once OA publications are beginning to appear readers "lower the threshold of effort they are willing to expend to retrieve documents that present *any* barriers to access. This indicates both a "push" away from print and a "pull" toward open access, which may strengthen the association between open access and research impact" (Antelman 2004: 377).

Notwithstanding all this, OA as it is managed today still presents serious shortcomings: "even if publishing in an open-access journal were generally associated with a 10% boost in citations, it is not clear that authors in economics and business would be willing to pay several thousand dollars for this benefit, at least in lieu of subsidies" (McCabe and Snyder: 2013: 31) referring to the OA models often adopted by commercial publishers. In many cases, national funding bodies require research results to be published in an OA environment, and due to indirect assessment – a model very often used for the evaluation of personal careers – with the ranking of journals and pub-lishers dictating where to publish (mostly commercial publishers and sub-

scription-based journals), and, thus, forcing upon researchers a rather expensive publication option, "authors simply have to go for the expensive Open Access strategy (aptly called 'Gold Open Access')" (Blommaert 2014: 3), thereby supporting a barefaced "robber economy" as a "no- risk enterprise in its most extreme shape" (Blommaert 2014: 4). If a researcher does not comply with this approach, insisting on his freedom of choosing other publication options, this often results in a lack of prestige when his/her articles or books are published in journals or with publishers that are not listed in the rankings.

Along with top ranking goes visibility of articles in a discipline, and, conversely, research results published in journals or with publishers which are not listed in the rankings may not be immediately appreciated by colleagues and fellow researchers. However, there are quite a few OA repositories and search platforms available today where OA publications can be searched for on the basis of their metadata, and downloaded:

- the OAIster Database (oaister.worldcat.org) with records of digital resources from open-archive collections worldwide;
- the Directory of Open Access Journals DOAJ (doaj.org) with more than 600 searchable journals;
- The Directory of Open Access Repositories – OpenDOAR (opendoar.org), a directory of academic open access repositories;
- BioMed Central (biomedcentral.com), Open Access journals covering all areas of Biology and Medicine;
- Public Library of Science (PloS) (plos.org), a nonprofit scientific and medical publishing venture using the Creative Commons Attribution License;
- PLEIADI Portal for the Italian Electronic Literature in Open and Institutional Archives (openarchives.it/pleiadi/);
- OAPEN Open Access Publishing in European Networks (oapen.org), an online library and publication platform;
- SHERPA/RoMEO, a database about publisher copyright policies & self-archiving options.

Openness in publishing and the institution of freely accessible publication archives even seem to promote the international ranking of universities as empirical studies show (Olsbo 2013); I will come back to the problems of evaluation and assessment of research in more detail below.

From the viewpoint of authors, scholars or researchers the positive aspects of OA clearly prevail: OA brings greater impact, dissemination of research results is faster, it enables better management and assessment of

research, and provides new opportunities for linking and online text-mining, as well as a degree of productive collaboration otherwise not possible.

Coming back to TS, a look at the relevant journals and their publishing policies seems to suggest that OA journals are on the rise. There are several listings of relevant journals in TS, amongst others:

- RETI (RETI n.d.): Revistes dels Estudis de Traducció et Interpretació of the Autonomus University of Barcelona lists a total of 421 titles with many journals from neighboring disciplines such as linguistics and literature, out of which 161 (38 %) are found to be OA.
- Another list of 55 journals publishing TS research, published on Academia.edu by James Hadley, reports 20 OA titles or 36% electronically available as PDF files free of charge and without any subscription fee.
- The European Society for TS (EST) has a draft listing of 125 journals, 57 of which are found to be OA (46%), 5 partly (4%), 3 limited (2%), 2 first issue only (2%) and 50 subscription-based (40%), 8 not declared (6%).
- The recent list of active Journals in TS by Franco Aixelá/Rovira-Esteva (2015) in the special issue of Perspectives sees a majority of OA titles, 58 or 52% against 54 or 48% with toll access, out of a total of 112 journals.

Not taking into account the different inclusion criteria depending on categorization and discipline boundaries, the average ratio of OA journals in these lists is a hefty 43%, a high percentage, also confirmed by a study for the European Commission which found that "18% of biology papers published in 2008-11 were open access from the start, and said that 57% could be read for free in some form, somewhere on the Internet, by April 2013" (Noorden 2014: 128). In addition, the OA options for the publication of monographs and edited volumes, in TS more important than journals (Franco Aixelá and Rovira-Esteva 2015: 270; AQU Workshop 2010: 7), with big publishing houses are increasing, even if many of them are offering OA only on a very expensive basis. Small publishing enterprises by local universities seem to be the best option at this time as their OA price policies are much more accessible to constantly under-funded researchers.

Today, OA has ceased to be a rather strange, or a niche publishing option, and already begins to rival traditional publishing methods. Seen from the viewpoint of researchers and put in more ideological terms, it boils down to the question: Do I want my ideas and research results to be sold by commercial companies with the respective financial burden on potential

readers, or do I want them to be open and accessible to as many readers as possible?

2 Social Media for Researchers

New media present researchers with new and totally independent publication options, each of which with specific advantages and disadvantages, as well as a varying degree of openness. Scholars may have personal websites where articles, studies and monographs can be made accessible after their publication in journals or books if copyright contracts allow them to do so – a method called self-archiving – or even original work published for the first time. The problem with this form of independent publishing is that it will be difficult or nearly impossible for authors to reach a clearly defined target audience, usually fellow researchers from the same discipline or scholars from wider neighboring subject fields. Though self-archiving facilitates free access to publications, it does nothing to support collaboration and communication between scholars.

Social media platforms for scholars try to remedy this by devising convenient collaborative websites which allow scholars to share their works, reach the intended audience and get feedback at the same time, they enable social interaction. While such tools are already very popular for general purposes on the Internet (Facebook, LinkedIn, Twitter), for photo sharing (Flickr, Instagram), for Video sharing (YouTube), etc. they are gaining popularity in academia as well, either as a substitute for self-archiving, as a secondary publication method, or simply as a place to discuss research results and ideas: "such sharing tools are, in effect, perhaps the most 'ecological' tool available at present" (Blommaert 2014: 11). Online community resources for scholars and scientists from many disciplines give their "members a place to create profile pages, share papers, track views and downloads, and discuss research" (Noorden 2014: 126). The most prominent examples (Noorden 2014) are briefly discussed here from the perspective of their openness.

2.1 Google Scholar

Google Scholar is a specialized tool to search for scholarly literature. It allows researchers to explore related works, citations, authors, publications, and proposes links to complete documents. Citations of individual publications can be checked to see how often a paper has been cited, who cited the publication in which document and whether the document is freely available.

In addition, Google scholar offers the possibility to create a kind of homepage for each researcher, called the public author profile, that incorporates

his/her publications and a citation analysis. The number of citations is indicated for each individual publication, as well as for the researcher in total, and compiled into the h-index (see below).

For researchers, Google Scholar represents a very powerful tool that reveals relevant links between publications and authors, and offers one of the most comprehensive citation analyses. Critics (Fell 2010) point out that the algorithms used by GS are not open or documented so that metrics cannot be verified. Citation analysis and scholarly metrics will be dealt with in the next chapter.

2.2 ResearchGate

ResearchGate is more focused on social interaction between scholars and restricts membership to academic researchers. Each member has a public profile with a list of publications, a synopsis of new publications in the field of research, a page with research questions regarding the specific discipline, as well as a scholarly metrics index, the RG-Score. This RG Score constitutes a rather unique index based on a proprietary design and computation basis. It seems to include the geographically and culturally very biased Thomson Reuters Web of Knowledge (WoK) database, on the one hand, as well as the researcher's social engagement on the platform, on the other hand: "anything researchers contribute to the network becomes a factor in their RG Score" (Tausch n.d.: 2). The RG Score changes on the basis of the scholars' involvement in the platform, independently of his/her publications, and is, thus, not well suited as a research assessment criterion: "We simply suggest to the ResearchGate decision makers to dump it into the dustbin of scientific errors and useless concepts, for good and forever" (Tausch n.d.: 3).

Overall, researchers seem to have reservations towards ResearchGate and their 'annoying policies' (Noorden 2014: 127), a geneticist, for example, is cited as saying "I've met basically no academics in my field with a favorable view of ResearchGate" (Noorden 2014: 126).

2.3 Academia.edu

Academia.edu is another popular social networking site for academics; according to their website "23,166,542 academics have signed up to Academia.edu, adding 6,167,754 papers" (July 2015). The site combines the feature of a publication archive integrating different document types with social networking capabilities, such as profiles, news feeds, recommendations, and the ability to follow individuals and subject fields or topics. The makers of Academia.edu stress their commitment to the principles of open science and open access.

2.4 ORCID

ORCID was conceived as an "open, non-profit, community-based effort to provide a registry of unique researcher identifiers and a transparent method of linking research activities and outputs to these identifiers" (ORCID website) to avoid misidentification and author ambiguity problems. By becoming a member and getting the ORCID ID code, each scholar can enter basic personal information and affiliation, as well as a list of publications. ORCID basically, represents a searchable database of researchers, and is recommended by the SPRU (2015) report to be the "preferred system of unique identifiers" for the UK research system.

2.5 ResearcherID

More or less the same functionality is offered by ResearcherID which is part of Thomson Reuters and integrates into their Web of Science database. It is a free tool by a commercial provider.

3 Research Evaluation

Open Access and new academic publishing and communication platforms lead to more openness with regard to potential readership, and more transparency in publishing. The OA citation effect gives researchers a clear advantage as to when, and how often their publications are read and cited by fellow scholars. While this may translate into a better reputation and a higher self-esteem it is by no means a matter of course that it has the same positive impact on assessment procedures for careers and tenures. Here, we need to discuss the degree of openness and transparency of the different models of research evaluation which are of overall importance for researchers who still need to secure their career or livelihood.

Evaluation may be performed by direct or indirect research quality assessment (Rovira-Esteva and Orero 2012: 270), where a direct approach evaluates the works of an individual scholar or research group by looking at the quality, relevance, citation rate, or impact factor of his/her/their publications, and an indirect approach evaluates the works of an individual scholar or research group by looking at the scientific performance (quality/relevance/citation rate/impact factor) of the journals, publishers, series where his/her/their works were published. The first can be more intricate and difficult while the second, it is argued, saves time by relying on the assessment of an already done peer-review and quality assessment of journals or publishers.

1. In both cases a variety of quantitative and qualitative metrics are used to measure productivity outcomes and impact of scholars, journals

and publishers, usually a combination of a quantitative analysis of publications – "authors, publication date, publication type, journal, publisher, etc., and statistical analyses in order to explain the growth (or decrease) of publication rates, the origin and evolution of disciplines, publication policy, interdisciplinarity, etc." (Grbić and Pöllabauer 2008: 5) –, a citation analyses by counting the citations of publications or journals to determine the impact on the discipline with the help of citation indexes and journal rankings, or a content analysis on publication data by measuring the occurrence and/or co-occurrence of certain keywords or subject classification categories in order to reveal trends regarding issues covered.

While counting publications seems to be sufficiently transparent, citation analysis is rather controversial. Basically, there are three ways in which citation analysis can be applied:

- to an individual article (how often it was cited);
- to an author (total citations, or average citation count per article);
- to a journal (average citation count for the articles in the journal), called the Journal Impact Factor (JIF).

To assess the impact, various calculations are done on the citation numbers and expressed in so-called impact factors. The most common is the h-index which "is a measure to quantify the cumulative impact of the publica-tions of a scholar or research community by looking at the number of times those works have been cited" (Grbić and Pöllabauer 2008), a research community (or scholar) with an index of 'H' has published 'H' papers, each of which has been cited at least 'H' times: "the higher the h-index, the more influential is the research community" (Xiangdong 2015: 185). Variations of the h-index such as the contemporary h-index or the individual h-index try to accommodate different parameters such as the number of authors per publi-cations into the calculus. The g-index complements the h-index by calculating the average citation rate of all publications of an author, also taking into account full citation numbers of very highly cited papers. A well documented tool which calculates H, G, and other indexes by using Google Scholar results is Harzing's Publish or Perish software (Harzing 2007).

While these data certainly provide an insight into the research impact of individual authors they should always be interpreted cautiously: different disciplines have divergent citation patterns or publication practices, such as the preference for book publications in humanities. Moreover, a citation may not always mean approval or recognition: the reason for citing a specific work could also be refusal or rejection, and the collection of citations may not be exhaustive as bibliographic databases tend to be work in progress.

The most used databases for citation analysis are two commercial applications, the Web of Science by Thomson Reuters with their Arts and Humanities Citation Index AHCI and the Scopus database by Elseviers, and the freely accessible Google Scholar database. While the completeness and coverage of publications of the Web of Science has been criticized heavily since it "may provide a substantial underestimation of an individual academic's actual citation impact" (Harzing and van der Wal 2008: 62), the problems of applying the two commercial indexes to the humanities in general – "the Social sciences, Arts and Humanities, and engineering in particular seem to benefit from Google Scholar's better coverage of (citations in) books, conference proceedings and a wider range of journals" (Harzings PoP website) – and TS in particular, have been emphasized repeatedly. Franco Aixelá and Rovira-Esteva (2015: 269) make clear that Google Scholar and Bitra, a specialized bibliographic database, are far more efficient in providing citations for articles in the subject field of TS than WoS/AHCI or Scopus; the latter do not treat TS as an autonomous discipline: "bibliometric tools such as BITRA or Google Scholar are beginning to provide a clearer picture of the impact of research in TS" (Franco Aixelá and Rovira-Esteva 2015: 277);

"Google Scholar results, even if it's not an index and data is mechanically gathered, throw a more objective and thorough results than the established and more valued indexes – with the added value of being free of access" (Rovira-Esteva and Orero 2012: 271).

Openness as free access also means the reproducibility of assessments, and, thus, more transparency:

"Google Scholar provides an avenue for more transparency in tenure reviews, funding and other science policy issues, as it allows citation counts, and analyses based thereon, to be performed and duplicated by anyone" (Harzing 2008).

But free access alone is not enough for complete openness, the underlying data and algorithms have to be open and verifiable as well (SPRU 2015: 6): this seems not to be the case with the Web of Science, Scopus, and even Google Scholar. Still, citation analysis of articles and individual scholars constitute a transparent and verifiable method of assessment: "article-level citation metrics, for instance, might be useful indicators of academic impact, as long as they are interpreted in the light of disciplinary norms and with due regard to their limitations" (SPRU 2015 recommendation n°4). Indirect assessment, in contrast, rates research work on the basis of where it has been published, using ratings or classifications of journals and publishers, thus, judging "our science by its wrapping rather than by its contents" (Seglen 1997: 501).

Indirect assessment should, therefore, generally be rejected: "Journal-level metrics, such as the JIF, should not be used" (SPRU 2015 recommendation 4), and "do not use journal-based metrics, such as Journal Impact Factors, as a surrogate measure of the quality of individual research articles, to assess an individual scientist's contributions, or in hiring, promotion, or funding decisions" (San Francisco Declaration on Research Assessment DORA, recommendation 1). The reasons for this rejection were appropriately summarized by Seglen (1997: 498):

- The JIF "conceals the difference in article citation rates (articles in the most cited half of articles in a journal are cited 10 times as often as the least cited half)
- Journals' impact factors are determined by technicalities unrelated to the scientific quality of their articles
- Journals' impact factors depend on the research field: high impact factors are likely in journals covering large areas of basic research with a rapidly expanding but short lived literature that use many references per article
- Article citation rates determine the journal impact factor, not vice versa" (Seglen 1997: 498)

These arguments are shared by other scholars as well: Antelman (2004), for example, states with regard to the difference in article citation rates that "the high standard deviations of these samples bear this out and point to the value of new citation measures [...] Open-access articles make these new, more meaningful measures of research impact possible" (Antelman 2004: 380). The JIF should be restricted to the evaluation of journals and, in no case be extended to the assessment of an individual's work since

"the quality, reputation and impact of journals are therefore not achievements of the journals and their publishers: they are overwhelmingly achieved by the academic community that furnishes top-quality materials to them. After all, it's not journals that are cited but articles" (Blommaert 2014: 2).

Leaving aside arguments of a more general nature, indirect assessment through the JIF or other citation indexes is even more questionable when the humanities or, more specifically, TS are concerned. The common indexes are not suited for the humanities "because of their unsatisfactory coverage of European humanities research" (Franco Aixelá and Rovira-Esteva 2015: 268), proven by practical verification: "of more than 100 TS journals throughout the world (including both English and non-English TS journals), only 13 are indexed in the SSCI (Social Sciences Citation Index) or AHCI (Arts & Humanities Citation Index) databases" (Xiangdong 2015: 184). This leads to a rather weak ranking of publications in TS. Even those listed are treated rather poorly in comparison to larger disciplines: "Impact Factors [...] of TS journals

are low compared with other Linguistics journals" (Xiangdong 2015: 184), with negative effects for researchers: "this means TS scholars would be put in a disadvantaged position when being assessed against the same research assessment policy to decide their assignment, research ranking, promotion, and research funding, compared with Linguistics scholars" (Xiangdong 2015: 184).

To sum up, openness in assessment can only be achieved if individual scholars and research groups are evaluated directly, without recurring to journal impact factors. On the way "to a more open, accountable and outward-facing research system" (SPRU 2015: 5), impact factors and numbers in general should better be avoided and supplanted by the term 'indicators' when the work of individual scholars is evaluated (SPRU 2015 recommendations). The Independent Review of the Role of Metrics in Research Assessment and Management (SPRU 2015) defines "responsible metrics" according to five parameters:

"Robustness: basing metrics on the best possible data in terms of accuracy and scope; Humility: recognising that quantitative evaluation should support – but not supplant – qualitative, expert assessment; Transparency: keeping data collection and analytical processes open and transparent, so that those being evaluated can test and verify the results; Diversity: accounting for variation by field, and using a variety of indicators to support diversity across the research system; Reflexivity: recognising systemic and potential effects of indicators and updating them in response" (SPRU 2015: 7).

Implementing the guidelines and applying these principles in practice would guarantee more openness in evaluation procedures and research assessment.

4 Conclusions

The more scholars accept and adopt openness in their work, the more collaboration between researchers will take place, the faster research work will be read and processed, and the fairer assessment procedures will be. In summary, the advantages of open scholarship may be outlined schematically in the following diagram where the three areas of literature search, open publishing, and research assessment each generate specific advantages amplified through interaction with each other:
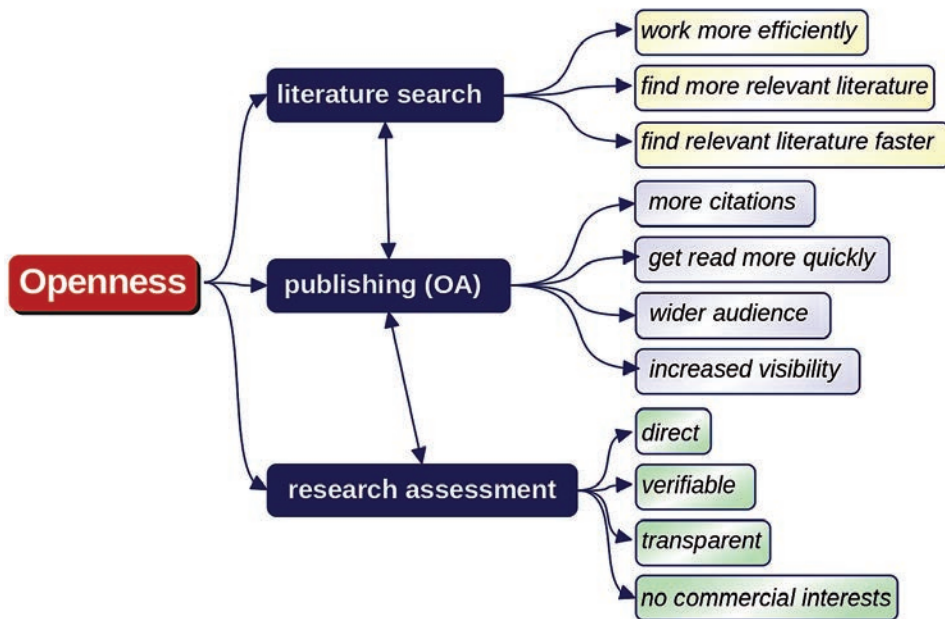
Figure 1: Advantages of openness.

A discipline can only gain from such an accelerated pace and transparent procedures, and, more importantly, isolated approaches and closed branches of theory will be avoided. This is especially important for TS where openness can help overcome ignorance and disregard of important literature as well as fragmentation of the discipline into mutually ignored schools of thought.

References

Antelman, K. (2004) Do Open-Access Articles Have a Greater Research Impact? Coll. res. libr. 9/65, 372-382; doi:10.5860/crl.65.5.372. Available at: http://crl.acrl.org/content/65/5/372.full.pdf+html [Accessed 3 August 2015].

AQU Workshop (2010) Research Assessment in the Humanities and Social Sciences. Eleventh Workshop – AQU Catalunya with the Catalan Universities, University of Barcelona, 28-29 January 2010. Available at http://www.aqu.cat/doc/doc_86235770_1.pdf [Accessed 3 August 2015].

Blommaert, J. (2014) The Power of Free: In search of democratic academic publishing strategies. Tilburg Papers in Cultural Studies. *Tilburg: Tilburg University.* Paper 114.

Davis, P. M., Lewenstein, B.V., Simon, D.H., Booth, J.G. and Connolly, M.J. (2008) Open access publishing, article downloads, and citations: randomised controlled trial. BMJ. BMJ Publishing Group Ltd. 337. Available at: http://www.bmj.com/content/337/bmj.a568 [Accessed 3 August 2015].

Davis, P. M. (2011) Open access, readership, citations: a randomized controlled trial of scientific journal publishing. The FASEB Journal. 25(7), 2129-2134. Available at: http://www.fasebj.org/content/25/7/2129.full.pdf [Accessed 2 August 2015].

Doty, R. C. (2013) Tenure-Track Science Faculty and the 'Open Access Citation Effect'. *Journal of Librarianship and Scholarly Communication* 1(3):eP1052. Available at: http://dx.doi.org/10.7710/2162-3309.1052 [Accessed 2 August 2015].

European Science Foundation ESF, Humanities unit (2009) Increasing visibility for a multifaceted humanities research in Europe – the ERIH approach. Available at: http://docslide.net/documents/erihpresentation2009.html [Accessed 3 August 2015].

Eysenbach, G. (2006) Citation Advantage of Open Access Articles, PLoS Biology 4(5): e157. Available at: http://journals.plos.org/plosbiology/article?id=10.1371/journal.pbio.0040157 [Accessed 26 June 2015].

Fell, C. (2010) Publish or Perish und Google Scholar – ein Segen? Leibniz-Zentrum für Psychologische Information und Dokumentation (ZPID), Trier. Available at: http://www.zpid.de/pub/research/2010_Fell_Publish-or-Perish.pdf [Accessed 4 August 2015].

Franco Aixelá, J. and Rovira-Esteva, S. (2015) Publishing and impact criteria, and their bearing on Translation Studies: In search of comparability. Perspectives: Studies in Translatology, Special Issue: Bibliometric and Bibliographical Research in Translation Studies, 23(2), 265-283.

Gentzler, E. (2014) Translation Studies: Pre-Discipline, Discipline, Interdiscipline, and Post-Discipline. International Journal of Society, Culture & Language, 2(2), 13-24. Available at: http://ijscl.net/pdf_5620_fdde5469d71359e7bb41dcee95329e13.html [Accessed 2 August 2015].

Gile, D. (2015) Analyzing Translation studies with scientometric data: from CIRIN to citation analysis. Perspectives: Studies in Translatology, Special Issue: Bibliometric and Bibliographical Research in Translation Studies, 23(2), 240-248.

Grbić, N. and Pöllabauer, S. (2008) To count or not to count: Scientometrics as a methodological tool for investigating research on translation and interpreting. Translation and Interpreting Studies 3. Amsterdam: John Benjamins Publishing, 87-146.

Hadley, J. (2014) List of Journals Publishing Translation Studies Research. Available at: https://www.academia.edu/11919672/List_of_Journals_Publishing_Translation_Studies_Research [Accessed 2 August 2015].

Harnad, S. and Brody T. (2004) Comparing the impact of open access (OA) vs. non-OA articles in the same journals. D-Lib Magazine 10(6). Available at: http://eprints.soton.ac.uk/ 260207/1/06harnad.html [Accessed 2 August 2015].

Harzing, A. W. (2007) Publish or Perish. Available at: http://www.harzing.com/pop.htm [Accessed 2 August 2015].

Harzing, A. W. (2008) Google Scholar – a new data source for citation analysis. Available at: http://www.harzing.com/pop_gs.htm [Accessed 2 August 2015].

Harzing, A. W. and van der Wal, R. (2008) Google Scholar: the democratization of citation analysis? Ethics in Science and Environmental Politics, Vol 8, 61-73. Available at: http://www.int-res.com/articles/esep2008/8/e008p061.pdf [Accessed 3 August 2015].

Lawrence S. (2001) Free online availability substantially increases a paper's impact. Nature 411, 521. Available at: http://www.nature.com/nature/journal/v411/n6837/full/411521a0.html [Accessed 3 August 2015].

McCabe, M. J. and Snyder, C. M. (2013) Does Online Availability Increase Citations? Theory and Evidence from a Panel of Economics and Business Journals (March 14, 2013). Available at: http://ssrn.com/abstract=1746243 [Accessed 3 August 2015].

Noorden, R. van (2014) Scientists and the Social Network. Nature, vol. 512, 126-129. Available at: http://www.nature.com/polopoly_fs/1.15711!/menu/main/topColumns/topLeftColumn/ pdf/512126a.pdf [Accessed 28 July 2015].

Norris, M., Oppenheim, C. and Rowland, F. (2008) The citation advantage of open-access articles. Journal of the American Society for Information Science and Technology. Wiley Subscription Services, Inc., A Wiley Company. 59, 1963-1972. Available at: http://onlinelibrary.wiley.com/doi/10.1002/asi.20898/abstract;jsessionid=A4CBF131B9059A8E8748907EC9C883C0.f04t02 [accessed 3 August 2015].

Olsbo, P. (2013) Does openness and open access policy relate to the success of universities? Information Services and Use, 33 (2), 87-91. doi:10.3233/ISU-130707 Available at: http://elpub.scix.net/cgi-bin/works/Show?_id=110_elpub2013 [Accessed 3 August 2015].

RETI (n.d.) Journals of Translation and Interpreting Studies. University Library of Barcelona. Available at: http://www.bib.uab.cat/human/acreditacions/planes/publiques/revistes/ revistescercaetieng.php [Accessed 3 August 2015].

Rovira-Esteva, S. and Orero, P. (2012) Evaluating quality and excellence in translation studies research: Publish or perish, the Spanish way. Babel 58(3), 264-288.

Rovira-Esteva, S., Orero, P. and Franco Aixelá, J. (2015) Bibliometric and bibliographical research in Translation Studies, Perspectives: Studies in Translatology, Special Issue: Bibliometric and Bibliographical Research in Translation Studies, 23(2), 159-160.

Seglen, P. O. (1997) Why the Impact Factor of Journals Should Not Be Used for Evaluating Research. British Medical Journal 314 (Feb. 1997), 498–502. Available at: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2126010/pdf/9056804.pdf [Accessed 3 August 2015].

Smith Rumsey, A. (2013) New-Model Scholarly Communication: Road Map for Change. Scholarly Communication Institute Reports 9, 2004-2011. University of Virginia Library, 157-188. Available at: http://www.uvasci.org/institutes-2003-2011/SCI-9-Road-Map-for-Change.pdf [Accessed 3 August 2015].

SPRU – Science Policy Research Unit (2015) The Metric Tide: Report of the Independent Review of the Role of Metrics in Research Assessment and Management. Executive Summary. Available at: http://www.hefce.ac.uk/media/HEFCE,2014/Content/Pubs/Independentresearch/2015/The,Metric,Tide/2015_metric_tide_executive_summary_and_recommendations.pdf [Accessed 3 August 2015].

Tausch, A. (n.d.) Researchgate, RG-Scores, or a true Research Gate to Global Research? On the limits of the RG factor and some scientometric evidence on how thecurrent RG score system discriminates against economic and social sciences and against the

developing countries. Available at: https://www.academia.edu/2460163/ Researchgate_
RG-Scores_or_a_true_Research_Gate_to_Global_Research_On_the_limits_of_the_
RG_factor_and_some_scientometric_evidence_on_how_the_current_RG_score_syste
m_discriminates_even_against_Nobel_Laureates_in_economics_and_against_the_dev
eloping_countries [Accessed 3 August 2015].

White, E. (2015) On Success and Working Openly in Science. OpenCon Community
Webcasts. Available at: http://figshare.com/articles/On_success_and_working_ openly_
in_ science/ 1476243 [Accessed 3 August 2015].

Willinsky, J. (2006) Access Principle: The Case for Open Access to Research and
Scholarship. Cambridge: MIT press. Available at: http://mitpress.mit.edu/sites/default/
files/titles/content/9780262512664_Download_the_full_text.pdf [Accessed 3 August
2015].

Willinsky, J. (2010) Open access and academic reputation. Annals of Library and
Information Studies. Cambridge: MIT Press. 57, 296-302.

Xiangdong, L. (2015) International visibility of mainland China Translation Studies
community: A scientometric study. Perspectives: Studies in Translatology, Special
Issue: Bibliometric and Bibliographical Research in Translation Studies, 23(2), 183-204.

Further Literature and Useful Readings

Abaitua Odriozola, J. K. (2001) Memorias de traducción en TMX compartidas por Internet (TMX-based translation memories shared on the Internet). *Tradumatica* 1. Available at: http://www.fti.uab.es/tradumatica/revista/num0/articles/jabaitua/art.htm.
Keywords: open and collaborative translation, open tools.

Alabau, Vicent Bonk, R., Buck, C., Carl, M., Casacuberta, F., Garcia-Martinez, M., Gonzalez, J., Koehn, P. Leiva, L. Mesa-Lao, B. et al. (2013) CASMACAT: An open source workbench for advanced computer aided translation. *The Prague Bulletin of Mathematical Linguistics* 100, 101-112.
Keywords: open tools.

Alonso Jiménez, E. (2015) Analysing the use and perception of Wikipedia in the professional context of translation. *JosTrans* 23, 89-117.
Keywords: open and collaborative translation.

Anastasiou, D. and Gupta, R. (2011) Comparison of crowdsourcing translation with Machine Translation. *Journal of Information Science* 37 (6), 637-659.
Keywords: open and collaborative translation, MT.

Aparicio, S. (2001) Memòries de traducció d'accés públic: creació, gestió i ús. In Chabás, J.; Cases, M. & Gaser, R. (eds.) *Proceedings. First International Conference on Specialized Translation, Barcelona, March 2-4, 2000,* 161-166. Universitat Pompeu Fabra.
Keywords: open tools.

Arenas, A. G. (2010) Exploring Machine Translation on the Web. *Revista Tradumatica* 08/2010, 03.
Keywords: open and collaborative translation, MT.

Arjona Reina, L. (2012) *Translations in Libre Software.* Master Thesis, Madrid: Universidad Rey Juan Carlos.
Keywords: open tools

Arjona Reina, L., Robles, G. and González-Barahona, J. (2013) A Preliminary Analysis of Localization in Free Software: How Translations Are Performed. In Petrinja, E., Succi, G., El Ioini, N. and Sillitti, A. (eds.) *Open Source Software: Quality Verification,* 153-167. Heidelberg: Springer. Available at: http://link.springer.com/book/10.1007%2F978-3-642-38928-3.
Keywords: open tools.

Armentano-Oller, C., Corbí-Bellot, A. M., Forcada, M. L., Ginestí-Rosell, M., Montava Belda, M. A., Ortiz-Rojas, S., Pérez-Ortiz, J. A., Ramírez-Sánchez, G. and Sánchez-Martínez, F. (2007) Apertium, una plataforma de código abierto para el desarrollo de sistemas de traducción automática. In Rodríguez Galván, J. and Palomo Duarte, M. (eds.) *Proceedings of the FLOSS International Conference 2007,* 5-20. Servicio de Publicaciones de la Universidad de Cadiz. (ISBN: 978-84-9828-124-8.)
Keywords: open tools, MT.

Austermühl, F. (2011) On Clouds and Crowds: Current Developments in Translation Technology. *T21N* 09, 1-26.
Keywords: open and collaborative translation.

Babych, B., Hartley, A., Kageura, K., Thomas, M. and Utiyama, M. (2012) MNH-TT: a collaborative platform for translator training. In ASLIB (ed.) *Translating and the Computer 34, Vol. 34*. London: ASLIB.
Keywords: open and collaborative translation.

Bailey, D. (2012) Software Localization: Open Source as a Major Tool for Digital Multilingualism. In Le Crosnier, H. and Vannini, L. (eds.) *Net.lang. Towards the multilingual cyberspace,* 204-219. C&F éditions.
Keywords: open tools.

Baldwin, A. (2009) Linguas OS is dead. Long live linguas OS. Available at: http://linguasos.blogspot.com.es/2009/10/linguas-os-is-dead-long-live-linguas-os.html.

Bergmann, F. (2005) Open-Source Software and Localization. An Introduction to OSS and its impact on the language industry. *Multilingual Computing & Technology* 70 (16/2), 55-58. Available at: http://www.project-open.com/whitepapers/oss-l10n/.
Keywords: open tools.

Bey, Y., Boitet, C. and Kageura, K. (2006) The TRANSBey prototype: an online collaborative wiki-based cat environment for volunteer translators. In *LREC-2006: Fifth International Conference on Language Resources and Evaluation. Third International Workshop on Language Resources for Translation Work, Research & Training (LR4Trans-III),* 49-54.
Keywords: open tools, open and collaborative translation.

Bey, Youcef, Boitet, C. and Kageura, K. (2008) BEYTrans: A Wiki-based environment for helping online volunteer translators. In Yuste Rodrigo, E. (ed.) *Topics in Language Resources for Translation and Localisation,* 135-150. John Benjamins Publishing.
Keywords: open and collaborative translation, open tools.

Bodeux, E. and McKay, C. (2010) Free and Open Source Software for Translators. Available at: http://speakingoftranslation.com/listen/podcast-archives/.

Boitet, C., Bey, Y. and Kageura, K. (2005) Main research issues in building web services for mutualized, non-commercial translation. In *Proceedings of the 6th Symposium on Natural Language Processing,* 451-454.
Keywords: open and collaborative translation.

Bold, B. (2011) The power of fan communities: An overview of fansubbing in Brazil. *Traduçao em Revista* 11, 2-19.
Keywords: open and collaborative translation, fansubbing.

Briel, D. (2011) Les outils libres du traducteur, un écosystème à apprivoiser. *Traduire* 224, 38-49.
Keywords: open tools.

Brunette, L. and Désilets, A. (2008) Quality in collaborative translation and terminology. *Multilingual* 9, 55-58.
Keywords: open and collaborative translation.

Calvert, D. (2008) Wiki behind the firewall – Microscale online collaboration in a translation agency. In ASLIB (ed.) *Translating and the Computer* 30, 27-28. London: ASLIB*.
Keywords: open and collaborative translation.

Campos, P. and Ramón, J. (2008) OpenTrad. Plataforma de tradución automática de código aberto [OpenTrad. An open-source platform for machine translation]. In Diaz Fouces, O. and García González, M. (eds.) *Traducir (con) software libre,* 123-136. Granada: Comares.
Keywords: open tools, MT.

Cánovas, M. (2008) Dos ejemplos de aplicación del software libre an la docencia de la traducción. In Diaz Fouces, O. and García González, M. (eds.) *Traducir (con) software libre,* 193-210. Granada: Comares.
Keywords: open tools.

Cánovas, M. and Samson, R. (2008) Herramientas libres para la traducción en entorno MS Windows. In Diaz Fouces, O. and García González, M. (eds.) *Traducir (con) software libre,* 33-56. Granada: Comares.
Keywords: open tools.

Canovas, M. and Samson, R. (2011) Open source software in translator training. *Tradumàtica: tecnologies de la traducció – Traducció i software lliure* 0(9), 46-56.
Keywords: open tools.

Castineira, G. (2008) Interfaces web na tradución de proxectos comunitarios de software libre. In Diaz Fouces, O. and García González, M. (eds.) *Traducir (con) software libre,* 137-157. Granada: Comares.
Keywords: open tools.

Chrupalstroka, G. (2003) Perl Scripting in Translation Project Management. *Across Languages and Cultures* 4 (1), 109-132. Available at: http://www.ingentaconnect.com/content/akiado/alc/2003/00000004/00000001/art00006.
Keywords: open tools.

Cordeiro, G. (2011) El software libre en la caja de herramientas del traductor. *Tradumàtica: tecnologies de la traducció – Traducció i software lliure* 9, 101-107. Available at: http://revistes.uab.cat/tradumatica/issue/view/9.
Keywords: open tools.

Cronin, M. (2010) The Translation Crowd. *Revista Tradumàtica* 8, 1-7.
Keywords: open and collaborative translation.

Cronin, M. (2013) *Translation in the digital age*. London: Routledge. (ISBN: 9780415608596).
Keywords: open and collaborative translation.

de la Fuente, L.C. (2014) La motivación del crowdsourcing multilingüe en los medios sociales globales. Un estudio de caso: TED OTP. *Sendebar* 25, 197-218.
Keywords: open and collaborative translation.

DePalma, D. (2008) Industry Dreams of Open-Source TMS. *Global Watchtower Blog.* Available at: http://www.commonsenseadvisory.com/Default.aspx?Contenttype=ArticleDetAD&tabID=63&Aid=538&moduleId=391.
Keywords: open tools.

DePalma, D. (2010) Open-Source Tools Support Machine Translation. *Global Watchtower Blog.* Available at: http://www.commonsenseadvisory.com/Default.aspx?Contenttype=ArticleDetAD&tabID=63&Aid=709&moduleId=391.
Keywords: open tools, MT.

DePalma, D. and Kelly, N. (2008) Translation of, for and by the people: How user-translated content projects work in real-life. Common Sense Advisory.
Keywords: open and collaborative translation.

Désilets, A. (2007) Translation wikified: How will massive online collaboration impact the world of translation? In ASLIB (ed.) *Translating and the computer 29*. London: Aslib.
Keywords: open and collaborative translation.

Désilets, A. (2010) Collaborative Translation: technology, crowdsourcing, and the translator perspective. *Introduction to workshop at AMTA* 3.
Keywords: open and collaborative translation.

Désilets, A., Barrière, C. and Quirion, J. (2008) Making wikimedia resources more useful for translators. In *Proceedings WikiMania'07: The International WikiMedia Conference*.
Keywords: open and collaborative translation.

Désilets, A., Gonzalez, L., Paquet, S. and Stojanovic, M. (2006) Translation the Wiki way. In *Proceedings of the 2006 International Symposium on Wikis,* 19-31. ACM Press. Available at: https://www.opensym.org/ws2006/proceedings/.
Keywords: open and collaborative translation.

Désilets, A., Huberdeau, L., Laporte, M. and Quirion, J. (2009) Building a Collaborative Multilingual Terminology System. In ASLIB (ed.) *Translating and the Computer, Vol. 29*. London: ASLIB.
Keywords: open and collaborative translation.

Désilets, A. and van der Meer, J. (2011) Co-creating a repository of best-practices for collaborative translation. *Linguistica Antverpiensia* 10, 27-46. Available at: http://scholar.google.comhttps://lans-tts.ua.ac.be/index.php/LANS-TTS/article/view/276.
Keywords: open and collaborative translation.

Diaz Cintas, J. (2009) *New Trends in Audiovisual Translation.* Bristol: Multilingual Matters.
Keywords: fansubbing, open and collaborative translation

Díaz Cintas, J. and Muñoz Sánchez, P. (2006) Fansubs: Audiovisual Translation in an Amateur Environment. *Journal of Specialised Translation* 06(7), 37-52. Available at: http://www.jostrans.org/issue06/issue06_toc.php.
Keywords: fansubbing, open and collaborative translation.

Diaz Fouces, O. (2005) Software libre en la formación de traductores: entre elpragmatismo y la utopía. In Romana García, M. L. (ed.) *II AIETI. Actas del II Congreso Internacional de la Asociación Ibérica de Estudios de Traducción e Interpretación. Madrid, 9-11 de febrero de 2005,* 25-43. Asociación Ibérica de Estudios de Traducción e Interpretación (AIETI).
Keywords: open tools.

Diaz Fouces, O. (2008) Ferramentas livres para traduzir com GNU/Linux e Mac OS X. In Diaz Fouces, O. and García González, M. (ed.) *Traducir (con) software libre,* 57-73. Granada: Comares.
Keywords: open tools.

Diaz Fouces, O. (2010) LaTeX en la formación de traductores: ¿y por qué no?. *Redit. Revista Electrónica de Didáctica de la Traducción y la Interpretación* 4.
Keywords: open tools.

Diaz Fouces, O. (2011) ¿Merece la pena introducir el software libre en la formación de traductores profesionales? In Universitat de Vic (ed.) *Anais das XI Jornadas de Traducción y Lenguas Aplicadas – Congreso Internacional "Didáctica de las lenguas y la traducción en la enseñanza presencial y a distancia" CDROM Language and Translation Teaching in FacetoFace and Distance Learning (2011).* Facultat de Ciències Humanes, Traducció i Documentació de la Universitat de Vic.
Keywords: open tools.

Diaz Fouces, O. (2011) Editorial: el programari lliure com a objectiu i com a instrument per a la traducció. *Tradumàtica. Tecnologies de la Traducció* 9, 1-4. Available at: http://revistes.uab.cat/tradumatica/issue/view/9.
Keywords: open tools.

Diaz Fouces, O. (2012) La naturaleza de las habilidades tecnológicas en la formación de traductores y el papel del software libre. In Cánovas, M.; Delgar, G.; Keim, L.; Khan, S. and Pinyana, A. (eds.) *Challenges in Language and Translation Teaching in the WEB 2.0 Era,* 159-167. Granada: Comares.
Keywords: open tools.

Diaz Fouces, O. (2012) Un proyecto de formación de traductores basado en software libre. *Anuario Brasileño de Estudios Hispánicos* 22, 56-65. Available at: http://revistes.uab.cat/tradumatica/issue/view/9.
Keywords: open tools.

Diaz Fouces, O. and García González, M. (eds.) (2008) *Traducir (con) software libre. Granada:* Editorial Comares. (ISBN: 978-8498364873.).
Keywords: open tools.

Dombek, M. (2014) *A study into the motivations of internet users contributing to translation crowdsourcing: the case of Polish Facebook user-translators.* Unpublished doctoral disseration, School of Applied Language and Intercultural Studies, Dublin City University. Available at: http://doras.dcu.ie/19774/.
Keywords: open and collaborative translation.

Drugan, J. (2011) Translation ethics wikified: How far do professional codes of ethics and practice apply to non-professionally produced translation? *Linguistica Antverpiensia, New Series – Themes in Translation Studies* 10 (10), 111-130.
Keywords: open and collaborative translation.

Drugan, J. and Babych, B. (2010) Shared resources, shared values? Ethical implications of sharing translation resources. In *Proceedings of the Second Joint EM+/CNGL Workshop "Bringing MT to the User: Research on Integrating MT in the Translation Industry"(JEC+ 10)* , 3-9.
Keywords: open and collaborative translation.

Dwyer, T. (2012) Fansub Dreaming on ViKi: "Don't Just Watch But Help When You Are Free". *The Translator* 18 (2), 217-243.
Keywords: open and collaborative translation, fansubbing.

EC DG, European Commission (2012) Studies on translation and multilingualism: Crowdsourcing translation. Available at: http://bookshop.europa.eu/de/crowdsourcing-translation-pbHC3112733/.
Keywords: open and collaborative translation.

Esplá-Gomis, M. (2009) Bitextor, a free/open-source software to harvest translation memories from multilingual websites. In *Beyond Translation Memories Workshop (MT Summit XII)*.
Keywords: open tools.

Esplà-Gomis, M., Sánchez-Martínez, F. and Forcada, M.L. (2015) Using on-line available sources of bilingual information for word-level machine translation quality estimation. In *Proceedings of the 18th Annual Conference of the European Association for Machine Translation,* 19-26.
Keywords: open tools, MT.

Federmann, C. (2010) Appraise: An Open-Source Toolkit for Manual Phrase-Based Evaluation of Translations. In *Proceedings of the International Conference on Language Resources and Evaluation,* 1731-1734. European Language Resources Association (ELRA). Available at: http://www.mt-archive.info/LREC-2010-Federmann.pdf.
Keywords: open tools, MT.

Fernández Costales, A. (2012) Collaborative translation revisited: Exploring the rationale and the motivation for volunteer translation. *Forum* 10:1, 115-142.
Keywords: open and collaborative translation.

Fernández Costales, A. (2013) Crowdsourcing and Collaborative Translation: Mass Phenomena or Silent Threat to Translation Studies? *Hermeneus* 10, 85-110.
Keywords: open and collaborative translation.

Fernández García, J. R. (2003) La Traducción en el mundo del Software Libre: Análisis del estado de las herramientas lingüísticas, proyectos actuales y necesidades de la comunidad del software libre. Available at: http://es.tldp.org/Articulos/0000otras/doc-traduccion-libre/.
Keywords: open tools.

Fernández García, J. R. (2011) Algunes reflexions sobre la localització comunitària de programari lliure. *Tradumàtica: tecnologies de la traducció – Traducció i software lliure* 9, 12-34. Available at: http://revistes.uab.cat/tradumatica/issue/view/9.
Keywords: open tools.

Fernández García, J.R. (2006) La traducción de software libre. Cerrando el círculo. *Linux Magazine* 23, 73-76. Available at: http://people.ofset.org/jrfernandez/edu/n-c/traducc_5/index.html.
Keywords: open tools.

Fernández García, J.R. (2006) La traducción del software libre. Cambio de herramientas. *Linux Magazine* 22, 74-78.
Keywords: open tools.

Fernández García, J.R. (2006) La traducción del software libre. Oportunidad de colaborar. *Linux Magazine* 19, 76-80. Available at: http://revistes.uab.cat/tradumatica/issue/view/9.
Keywords: open tools.

Fernández Pintelos, M. J. (2011) Traducción automática y software libre en la formación de traductores. *Translation Directory* 2309. Available at: http://www.translationdirectory.com/articles/article2309.php.
Keywords: open standards and formats.

Fernández Pintelos, M. J. (2010) Software libre na universidade: o caso da Licenciatura de Tradución e Interpretación. *Viceversa. Revista Galega de Tradución* 16.
Keywords: open tools.

Flórez, S. (2013) *Tecnologías libres para la traducción y su evaluación.* Phd Unpublished doctoral disseration, Universitat Jaume I. Available at: http://traduccionymundolibre.com/ wiki/File:Florez-2013-tesis.pdf.
Keywords: open tools.

Flórez, S. and Alcina, A. (2011) Catálogo de software libre para la traducción. *Tradumàtica: tecnologies de la traducció - Traducció i software lliure* 0(9), 57-73.
Keywords: open tools.

Flórez, S. and Alcina, A. (2011) Free/Open-Source Software for the Translation Classroom. A Catalogue of Available Tools. *The Interpreter and Translator Trainer* 5(2), 325-357.
Keywords: open tools.

Folaron, D. (2010) Networking and volunteer translators. In Gambier, Y. and van Doorslaer, L. (eds.) *Handbook of Translation Studies,* 231-234. Amsterdam: John Benjamins.
Keywords: open and collaborative translation.

Forcada, M. L. (2006) Open source machine translation: an opportunity for minor languages. In *Proceedings of the Workshop "Strategies for developing machine translation for minority languages", LREC, Vol. 6,* 1-6. Available at: http://www.isca-speech.org/archive_open/saltmil/ SALTMIL2006_procs2006.pdf.
Keywords: open tools, MT.

Forcada, M. L., Ginestí-Rosell, M., Nordfalk, J., O'Regan, J., Ortiz-Rojas, S., Pérez-Ortiz, J. A., Sánchez-Martínez, F., Ramírez-Sánchez, G. and Tyers, F.M. (2011) Apertium: a free/open-source platform for rule-based machine translation. *Machine Translation* 25 (2), 127-144.
Keywords: open tools, MT.

Franco Aixelá, J. and Rovira-Esteva, S. (2015) Publishing and impact criteria, and their bearing on Translation Studies: In search of comparability. *Perspectives. Studies in Translatology* 23(2), 265-283. (Doi: 10.1080/0907676X.2014.972419.)
Keywords: open access.

Frimannsson, A. and Hogan, J. (2005) Adopting standards based XML file formats in open source localisation. *Localisation Focus–The Internation Journal of Localisation* 4, 9-23.
Keywords: open standards and formats, open tools.

García González, M. (2008) Free Software for Translators: Is the Market Ready for a Change? In Diaz Fouces, O. and García González, M. (eds.) *Traducir (con) software libre,* 9-31. Granada: Comares.
Keywords: open tools.

García González, M. (2013) Free and Open Source Software in Translator Education. The MINTRAD Project. *The International Journal for Translation & Interpreting Research* 5-2, 125-148.
Keywords: open tools.

García, E. (2012) Itzulpenak egiteko kode irekiko eta doako laguntzak [Open-code translation tools]. *Senez 43* 43, 203-218.
Keywords: open tools.

Garcia, I. (2009) Beyond translation memory: Computers and the professional translator. *The Journal of Specialised Translation* 12 (12), 199-214.
Keywords: open tools.

Gile, D. (2015) Analyzing Translation studies with scientometric data: from CIRIN to citation analysis. *Perspectives. Studies in Translatology* 23(2), 240-248. (Doi: 10.1080/0907676X.2014.972418.)
Keywords: open access.

Gomes de Oliveira, R. and Anastasiou, D. (2011) Comparison of SYSTRAN and Google Translate for English→Portuguese. *Tradumàtica: tecnologies de la traducció - Traducció i software lliure* 9, 118-136. Available at: http://revistes.uab.cat/ tradumatica/ issue/view/9.
Keywords: open tools, MT.

Gomis Parada, C. (2011) La localización de aplicaciones de software libre en el ámbito de la empresa. *Tradumàtica: tecnologies de la traducció – Traducció i software lliure* 9, 108-117. Available at: http://revistes.uab.cat/tradumatica/issue/view/9.
Keywords: open tools.

Goncharova, Y. and Lacour, P. (2011) TraduXio : nouvelle expérience en traduction littéraire. *Traduire* 225, online. (Doi: 10.4000/traduire.94.). Available at: http://traduire.revues.org/94.
Keywords: open and collaborative translation.

González, M. (2006) Traducción de software libre. Available at: http://www.asturlinux.org/ archivos/jornadas2006/ponencias/traduccion-mikel/traduccion.pdf.
Keywords: open tools.

Grbić, N. and Pöllabauer, S. (2008) To count or not to count: Scientometrics as a methodological tool for investigating research on translation and interpreting. *Translation and Interpreting Studies* 3 (1-2), 87-146.
Keywords: open access.

Guillardeau, S. (2009) *Freie Translation Memory Systeme für die Übersetzungspraxis: Ein kritischer Vergleich.* Unpublished master's thesis, Translationszentrum der Universität Wien.
Keywords: open tools.

Guyon, A. (2010) Grandeurs et misères de la traduction collaborative en ligne – The ups and downs of online collaborative translation. *L'Actualité langagière* 7(1), 33-36. Available at: http://publications.gc.ca/collections/collection_2010/tpsgc-pwgsc/S52-4-7-1.pdf.
Keywords: open and collaborative translation.

Hanes, W. F. (2012) Translating "real life": a case study of fan translation in the service of meme transmission. *In-Traduções Revista do Programa de Pós-Graduação em Estudos da Tradução da UFSC* 3 (4), 1-12.
Keywords: open and collaborative translation.

Hartley, T. (2009) Technology and translation. In Munday, J. (ed.) *The Routledge Companion to Translation Studies,* 106-127. Routledge.
Keywords: open tools.

Hernández Guerrero, M.J. (2014) La traducción de letras de canciones en la web de aficionados Lyrics Translate. com. *Babel* 60 (1), 91-108.
Keywords: open and collaborative translation.

Huberdeau, L.P., Paquet, S. and Désilets, A. (2008) The Cross-Lingual Wiki Engine: enabling collaboration across language barriers. In *WikiSym '08 Proceedings of the 4th International Symposium on Wikis,* 1-14. ACM. Available at: http://dl.acm.org/citation.cfm?id=1822276.
Keywords: open and collaborative translation.

Hyde, A. (2011) Open Translation Tools*. Floss Manuals*. Available at: http://en.flossmanuals.net/open-translation-tools/_booki/open-translation-tools/open-translation-tools.pdf.
Keywords: open tools.

Inose, H. (2012) Scanlation – What Fan Translators of Manga Learn in the Informal Learning Environment. In *International Symposium on Language and Communication: Research Trends and Challenges, IICS (Institute of Language and Communication Studies), Izmir University, 10th-13th June, Izmir (Turkey),* 73-84.
Keywords: open and collaborative translation.

Ivarsson, F. (2007) *Undertextning: Kvalitetsskillnader mellan professionella undertextare och amatörer (Quality comparison between professional and amateur subtitlers).* MA thesisUnpublished master's thesis, Linköping Univ.
Keywords: open and collaborative translation, fansubbing.

Jimenez-Crespo, M. A. (2013) Crowdsourcing, corpus use, and the search for translation naturalness: A comparable corpus study of Facebook and non-translated social networking sites. *Translation and Interpreting Studies* 8 (1), 23-49.
Keywords: open and collaborative translation.

Kageura, K., Abekawa, T., Utiyama, M., Sagara, M. and Sumita, E. (2011) Has translation gone online and collaborative? An experience from Minna no Hon'yaku. *Linguistica Antverpiensia* 10, 47-72.
Keywords: open and collaborative translation.

Karsch, B. I. (2014) Terminology work and crowdsourcing: Coming to terms with the crowd. In Kockaert, H.J. and Steurs, F. (eds.) *Handbook of Terminology*, Vol 1, 291-303. Amsterdam: John Benjamins Publishing Company.
Keywords: open and collaborative translation.

Kelly, N., Ray, R. and DePalma, D. (2011) From crawling to sprinting: Community translation goes mainstream. *Linguistica Antverpiensia* 10, 75-94.
Keywords: open and collaborative translation.

Klaus, C. (2014) *Translationsqualität und Crowdsourced Translation: Untertitelung und ihre Bewertung am Beispiel des audiovisuellen Mediums TEDTalk*, Vol. 68. Berlin: Frank & Timme.
Keywords: open and collaborative translation.

Kleijn, A. (2012) Open-Source-Software für Übersetzer. Available at: http://www.heise.de/open/artikel/Open-Source-fuer-Uebersetzer-1204029.html.
Keywords: open tools.

Koehn, P. (2009) A Web-Based Interactive Computer Aided Translation Tool. In *ACLDemos '09 Proceedings of the ACL-IJCNLP 2009 Software Demonstrations,* 17-20. Available at: http://www.aclweb.org/anthology/P/P09/P09-4005.pdf.
Keywords: open tools, MT.

Koehn, P., Hoang, H., Birch, A., Callison-Burch, C., Federico, M., Bertoldi, N., Cowan, B., Shen, W., Moran, C., Zens, R. at al. (2007) Moses: Open source toolkit for statistical machine translation. In *Proceedings of the 45th annual meeting of the ACL on interactive poster and demonstration sessions,* 177-180.
Keywords: open tools, MT.

Koehn, P. and Senellart, J. (2010) Convergence of translation memory and statistical machine translation. In *Proceedings of AMTA Workshop on MT Research and the Translation Industry,* 21-31.
Keywords: open tools, MT.

Lacour, P., Bénel, A., Eyraud, F., Freitas, A. and Zambon, D. (2010) TIC, collaboration et traduction : vers de nouveaux laboratoires numériques de translocalisation culturelle. *Meta* 55 (4), 674-692.
Keywords: open and collaborative translation.

Lacour, P., Freitas, A., Bénel, A., Eyraud, F. and Zambon, D. (2011) Translation and the New Digital Commons. Available at: http://lodel.irevues.inist.fr/tralogy/index.php?id=150.
Keywords: open and collaborative translation.

Lee, H. K. (2011) Participatory media fandom: A case study of anime fansubbing. *Media, Culture & Society* 33 (8), 1131-1147.
Keywords: open and collaborative translation, fansubbing.

Lesch, H. M. (2014) Vertaalpraktyke in die sosiale media:'n verbeterde vertaalteks vir'n virtuele gemeenskap?. *Tydskrif vir Geesteswetenskappe* 54 (1), 129-143.
Keywords: open and collaborative translation.

Lewis, D., Liu, Q., Finn, L., Hokamp, C., Sasaki, F. and Filip, D. (2014) Open, Web-based Internationalization and Localization Tools. *Translation Spaces* III, 99-132.
Keywords: open tools.

Lieske, C., McCormick, S. and Thurmair, G. (2001) The Open Lexicon Interchange Format (OLIF) Comes of Age. In Maegaard, B. (ed.) *Machine Translation in the Information Age. Machine Translation Summit VIII Proceedings,* 6. IAMT / European Association for Machine Translation. Available at: http://www.eamt.org/summitVIII/papers.html.
Keywords: MT, open standards and formats.

Losse, K. (2008) Achieving Quality in a Crowd-sourced Translation Environment . *Keynote lecture at the 13th Localisation Research Conference Localisation4All.* Dublin.
Keywords: open and collaborative translation.

Manuel Jerez, J., López Cortés, J. and Brander de la Iglesìa, M. (2004) Traducción e interpretación: Voluntariado y compromiso social. El compromiso social en traducción e interpretación: Una visión desde ECOS, traductores e intérpretes por la solidaridad [Translation and interpreting: Volunteering and social commitment. Social commitment in translation and interpreting. A vision from ECOS, translators and interpreters for solidarity]. *Puentes* 4, 65-72. Available at: http://www.ugr.es/%7Egreti/revista_ puente_ pdf.htm.
Keywords: open and collaborative translation.

Martínez-Gómez, A. (2015) Bibliometrics as a tool to map uncharted territory: A study on non-professional interpreting. *Perspectives. Studies in Translatology* 23(2), 205-222.
Keywords: open access, open and collaborative translation.

Mas, J. (2003) El software libre y las lenguas minoritarias: una oportunidad impagable. *Digithum* 5. Available at: http://www.uoc.edu/humfil/articles/esp/mas0303/ mas0303.html.
Keywords: open tools.

Massidda, S. (2015) *Audiovisual Translation in the Digital Age: The Italian Fansubbing Phenomenon.* London: Palgrave Macmillan.
Keywords: open and collaborative translation, fansubbing.

Mata Pastor, M. (2008) Formatos libres en traducción y localización. In Diaz Fouces, O. and García González, M. (eds.) *Traducir (con) software libre,* 75-122. Granada: Comares.
Keywords: open standards and formats.

Mayor, A., Alegria, I., De Ilarraza, A.D., Labaka, G., Lersundi, M. and Sarasola, K. (2011) Matxin, an open-source rule-based machine translation system for Basque. *Machine translation* 25 (1), 53-82.
Keywords: open tools, MT.

McDonough Dolmaya, J. (2011) A window into the professon: What translation blogs have to offer Translation Studies. *The Translator* 17(1), 77-104.
Keywords: open and collaborative translation.

McDonough Dolmaya, J. (2011) The ethics of crowdsourcing. *Linguistica Antverpiensia* 10, 97-110.
Keywords: open and collaborative translation.

McDonough Dolmaya, J. (2012) Analyzing the Crowdsourcing Model and Its Impact on Public Perceptions of Translation. *The Translator* 18:2, 167-191.
Keywords: open and collaborative translation.

McDonough Dolmaya, J. (2014) Revision history: Translation trends in Wikipedia. *Translation Studies* 8:1, 16-34.
Keywords: open and collaborative translation.

McKay, C. (2006) Free and Open Source Software for translators. *Panacea* Vol. VII, NBε 23. Junio, online. Available at: http://www.medtrad.org/panacea/IndiceGeneral/n 23_tribuna_McKay.pdf.
Keywords: open tools.

Melby, A. (2008) TBX-Basic Translation-Oriented Terminology Made Simple. *Tradumatica* 6.
Keywords: open standards and formats.

Morado Vázquez, L. and Wolff, F. (2011) Bringing industry standards to Open Source localisers: a case study of Virtaal. *Tradumàtica: tecnologies de la traducció – Traducció i software lliure* 9, 74-83. Available at: http://revistes.uab.cat/tradumatica/issue/view/9.
Keywords: open standards and formats.

Muñoz Sánchez, P. (2007) Romhacking: localización de videojuegos clásicos en un con-texto de aficionados. *Tradumatica* 5. Available at: http://revistes.uab.cat/tradumatica/ issue/view/
Keywords: open and collaborative translation.

Muñoz, J. M. and Vella Ramírez, M. (2010) Gestores de memorias de traducción de software libre. *Sendebar Revista de la Facultad de Traducción e Interpretación* 21 (21), 231-250.
Keywords: open tools.

Nord, C., Khoshsaligheh, M. and Ameri, S. (2015) Socio-Cultural and Technical Issues in Non-Expert Dubbing: A Case Study. *International Journal of Society, Culture & Language* 4, 1-16.
  Keywords: open and collaborative translation, fansubbing.

Norris, M., Oppenheim, C. and Rowland, F. (2008) The citation advantage of open-access articles. *Journal of the American Society for Information Science and Technology* 59 (12), 1963-1972. (Doi: 10.1002/asi.20898). Available at: http://dx.doi.org/10.1002/ asi.20898.
  Keywords: open access.

Notley, T., Salazar, J.F. and Crosby, A. (2015) Online video translation and subtitling: examining emerging practices and their implications for media activism in South East Asia. *Global Media Journal: Australian Edition* 9(1) (1).
  Keywords: open and collaborative translation.

O'Brien, S. (2011) Collaborative translation. In Gambier, Y. and van Doorslaer, L. (eds.) *Handbook of Translation Studies* 2, 17-20. Amsterdam: John Benjamins.
  Keywords: open and collaborative translation.

O'Brien, S. and Schäler, R. (2010) Next generation translation and localization: Users are taking charge. In ASLIB (ed.) *Translating and the Computer 32,* 18-19. London: ASLIB.
  Keywords: open and collaborative translation.

O'Hagan, M. (2008) Fan translation networks: an accidental translator training environment? In Kearns, J. (ed.) *Translator and interpreter training: Issues, methods and debates,* 158-183. London: continuum.
  Keywords: open and collaborative translation.

O'Hagan, M. (2009) Evolution of user-generated translation: Fansubs, translation hacking and crowdsourcing. *Journal of Internationalisation and Localisation* 1 (1), 94-121.
  Keywords: open and collaborative translation.

O'Hagan, M. (2011) Community Translation: Translation as a social activity and its possible consequences in the advent of Web 2.0 and beyond. *Linguistica Antverpiensia* 10, 1-10.
  Keywords: open and collaborative translation.

O'Hagan, M. (2012) From Fan Translation to Crowdsourcing: Consequences of Web 2.0 User Empowerment in Audiovisual Translation. *Approaches to Translation Studies* 36, 25-41.
  Keywords: open and collaborative translation, fansubbing.

O'Hagan, M. (2012) From Fan Translation to Crowdsourcing: Consequences of Web 2.0 User Empowerment in Audiovisual Translation. In Remael, A., Orero, P. and Carroll, M. (eds.) *Audiovisual Translation and Media Accessibility at the Crossroads. Media for All 3,* 25-41. Amsterdam: Rodopi.
  Keywords: open and collaborative translation.

O'Hagan, M. (2012) Translation as the new game in the digital era. *Translation Spaces* 1 (1), 123-141.
  Keywords: open and collaborative translation.

O'Hagan, M. (forthcoming). Massively Open Translation: Unpacking the relationship between technology and translation in the 21st century. *International Journal of Communication 8.*
  Keywords: open and collaborative translation.

Olohan, M. (2014) Why do you translate? Motivation to volunteer and TED translation. *Translation Studies* 7:1, 17-33. Available at: http://dx.doi.org/10.1080/14781700.2013.781952.
Keywords: open and collaborative translation.

Orrego-Carmona, D. (2014) Where is the audience? Testing the audience reception of non-professional subtitling. In Torres-Simon, E. and Orrego-Carmona, D. (eds.) *Translation Research Projects 5,* 77-92. Tarragona: Intercultural Studies Group. Available at: http://isg.urv.es/publicity/isg/publications/trp_5_2014/index.htm.
Keywords: open and collaborative translation, fansubbing.

Paquet, S., Désilets, A. and de Pedro, X. (2008) Babel wiki workshop: cross-language collaboration. In *WikiSym '08 Proceedings of the 4th International Symposium*. Available at: http://wikisym.org/ws2008/proceedings/3_workshops/302_BabelWiki.pdf.
Keywords: open and collaborative translation.

Peeters, J. (2011) *Traduction et communautés.* Artois: Presses Université.
Keywords: open and collaborative translation

Perea Sardón, J. I. (2010) *Revisión asistida por ordenador de traducciones. Aplicación práctica a la revisión del sistema operativo libre Ubuntu como ejemplo.* Phd Unpublished doctoral disseration, Universidad de Granada. Available at: http://opentranslation.es/recursos/PereaSardon2010.pdf.
Keywords: open tools, open and collaborative translation.

Perea Sardón, J. I. (2011) La revisió de les traduccions de programari lliure. *Tradumatica* 9, 35-45. Available at: http://revistes.uab.cat/tradumatica/issue/view/9.
Keywords: open tools, open and collaborative translation.

Pérez González, L. (2007) Fansubbing anime: Insights into the 'butterfly effect'of globalisation on audiovisual translation. *Perspectives* 14 (4), 260-277.
Keywords: open and collaborative translation, fansubbing.

Pérez-González, L. (2012) Co-creational subtitling in the digital media: Transformative and authorial practices. *International Journal of Cultural Studies* 16, 3-21.
Keywords: open and collaborative translation.

Pérez-González, L. (2013) Amateur subtitling as immaterial labour in digital media culture: An emerging paradigm of civic engagement. *Convergence: The International Journal of Research into New Media Technologies* 19(2), 1-19.
Keywords: open and collaborative translation, fansubbing.

Pérez-González, L. and Susam-Saraeva, Ş. (2012) Non-professionals translating and interpreting: Participatory and engaged perspectives. *The Translator* 18 (2), 149-165.
Keywords: open and collaborative translation.

Pérez, R. A. (2008) Software libre y/o gratuito de ayuda al traductor. In Union Latina (ed.) *Lenguas y dialogo intercultural en un mundo en globalizacion. Actas del Congreso Mundial de Traducción Especializada, Cuba dicembre 2008,* 391-394. Unión Latina.
Keywords: open tools.

Perrino, S. (2009) User-generated Translation: The future of translation in a Web 2.0 environment. *JostTrans The Journal of Specialised Translation* 12, 55-78.
Keywords: open and collaborative translation.

Popović, M. (2011) Hjerson: An open source tool for automatic error classification of machine translation output. *The Prague Bulletin of Mathematical Linguistics* 96, 59-67.
Keywords: open tools, MT.

Possamai, V. (2009) Catalogue of Free-Access Translation-Related Corpora. *Tradumatica* 7. Available at: http://www.fti.uab.cat/tradumatica/revista/num7/articles/09/09.pdf.
Keywords: open access, open and collaborative translation.

Prior, M. (2003) Close Windows. Open Doors. *Translation Journal* 7:1. Available at: http://translationjournal.net/journal/.
Keywords: open tools.

Prior, M. (2010) The open-source model. *ITI Bulletin* 1/2, 10. Available at: http://api.ning.com/files/sFo4PteLD*dcV54N6FflC6GrrA-BvwvCaM*W8cgbw7zzpzy3YrMkmQdgr8tw9AXWE36EDhqZhLs3iK*qpj5xkBBRVtApU86o/sample_issue_2010_01.pdf.
Keywords: open tools.

Prodan, D. I. (2008) Aportacions lingüístiques a sistemes oberts de traducció automàtica. El cas Apertium" [Linguistic contributions to open systems in machine translation. The case of Apertium]. In Navarro Domínguez, F., Vega Cernuda, M.A., Albaladejo Martínez, J. A., Gallego Hernández, D. and Tolosa Igualada, M. (eds.) *La traducción: balance del pasado y retos del futuro,* 201-210. Alicante: Departamento de Traducción e Interpretación de la Universidad de Alicante & Aguaclara.
Keywords: MT, open tools.

Proz.com (2014) Free Software for Translators: Free & Open Source Software for Translators. Available at: http://wiki.proz.com/wiki/index.php/Free_Software_for_Translators.
Keywords: open tools.

Pym, A. (2011) Democratizing translation technologies – the role of humanistic research. In *Luspio Translation Automation Conference, Vol. 5*. Available at: http://usuaris.tinet.cat/apym/on-line/research_methods/2011_rome_formatted.pdf.
Keywords: open tools.

Ramírez Polo, L. (2012) Software libre y software gratuito para la traducción. In Candel Mora, M. and Ortega Arjonilla, E. (eds.) *Tecnología, traducción y cultura,* 117-142. Valencia: Tirant Humanidades.
Keywords: open tools.

Ramírez-Sánchez, G., Sánchez-Martínez, F., Ortiz-Rojas, S., Pérez-Ortiz, J. A. and Forcada, M. (2006) Opentrad Apertium open-source machine translation system: an opportunity for business and research. In *Proceedings of Translating and the Computer 28 Conference*. (ISBN: 0-85142-483-X). Available at: http://www.dlsi.ua.es/~mlf/docum/ramirezsanchez06p.pdf.
Keywords: open tools, MT.

Ray, R. and Kelly, N. (2011) Crowdsourced translation: Best practices for implementation. Common Sense Advisory. Available at: https://www.commonsenseadvisory.com/Portals/_default/Knowledgebase/ArticleImages/110201_R_GL_Crowdsourcing_Preview.pdf.
Keywords: open and collaborative translation.

Sánchez-Cartagena, V.M., Sánchez-Martínez, F. and Pérez-Ortiz, J. A. (2012) An open-source toolkit for integrating shallow-transfer rules into phrase-based satistical machine translation. In *Proceedings of the Third International Workshop on Free/Open-Source Rule-Based Machine Translation,* 41-54.
Keywords: open tools, MT.

Sánchez-Martínez, F. and Forcada, M. (2011) Free/open-source machine translation: preface. *Machine Translation* 25 (2), 83-86.
Keywords: MT, open tools.

Sandrini, P. (2012) Tecnologia FLOSS per la traduzione: Disponibilità, applicazione e problematiche. *inTRAlinea* Special Issue: Specialized Translation II, no pages. Available at: http://www.intralinea.org/specials/article/1796.
Keywords: open tools.

Sandrini, P. (2012) Translationstechnologie im Curriculum der Übersetzerausbildung. In Malgorzewicz, A. and Zybatow, L. (eds.) *Sprachenvielfalt in der EU und Translation. Translationstheorie trifft Translationspraxis*, 107-120. Dresden-Wroclaw: Neisse. Available at: http://www.uibk.ac.at/downloads/trans/publik/transtech-wroclaw.pdf.
Keywords: open tools.

Sandrini, P. (2013) Open Translation Data – Die gesellschaftliche Funktion von Übersetzungsdaten. In Nord, B. and Mayer, F. (eds.) *Aus Tradition in die Zukunft. Perspektiven der Translationswissenschaft. Festschrift für Christiane Nord,* 27-37. Berlin: Frank & Timme. Available at: http://www.uibk.ac.at/downloads/trans/ publik/fsnord.pdf
Keywords: open standards and formats.

Sandrini, P. (2014) Open Translation Data: Neue Herausforderung oder Ersatz für Sprachkompetenz? In Ortner, H., Pfurtscheller, D., Rizzolli, M. and Wiesinger, A. (eds.) *Datenflut und Informationskanäle,* 185-194. Innsbruck: IUP. Available at: http://www.uibk.ac.at/downloads/trans/publik/medien-opendata.pdf.
Keywords: open standards and formats.

Sasikumar, M., Aparna, R., Naveen, K. and Rajendra Prasat M. (2005) *Free/Open Source Software – Guide to Localisation.* CDAC Centre for Development of Advanced Computing. Available at: http://www.eldis.org/go/topics/resource-guides/icts-for-development/open-development&id=20265&type=Document#.VX62eryYr3C.
Keywords: open tools

Simó Ferrer, R. M. (2005) Fansubs y scanlations: la influencia del aficionado en los criterios profesionales. *Puentes* 6, 27-44.
Keywords: open and collaborative translation, fansubbing.

TAUS (2011) Open Translation Platforms: Q&A PANEL [Vídeo on-line] Sta. Clara. Available at: http://www.youtube.com/watch?v=5uQ28kh0fXk &lr=1.
Keywords: open and collaborative translation.

Tyers, F. M., Sánchez-Martínez, F., Ortiz-Rojas, S. and Forcada, M.L. (2010) Free/Open-Source Resources in the Apertium Platform for Machine Translation Research and Development. *The Prague Bulletin of Mathematical Linguistics* 93 (93), 67-76.
Keywords: MT, open tools.

Wang, F. (2014) Similarities and Differences between Fansub Translation and Traditional Paper-based Translation. *Theory and Practice in Language Studies* 4 (9), 1904-1911.
Keywords: open and collaborative translation, fansubbing.

Wasala, Asanka, O'Keeffe, I. and Schäler. R. (2011) Towards an Open Source Localisation Orchestration Framework. *Tradumàtica: tecnologies de la traducció – Traducció i software lliure* 9, 74-83. Available at: http://revistes.uab.cat/tradumatica/issue/view/9.
Keywords: open tools.

Wilcock, S. (2013) *A comparative analysis of fansubbing and professional DVD subtitling.* Unpublished doctoral disseration, University of Johannesburg, ZA. Available at: https://ujdigispace.uj.ac.za/bitstream/handle/10210/8638/wilcock_2013.pdf?sequence=1.
Keywords: open and collaborative translation, fansubbing.

Xiangdong, L. (2015) International visibility of mainland China Translation Studies community: A scientometric study. *Perspectives. Studies in Translatology* 23(2), 183-204.
Keywords: open access.

Yan, Rui, Gao, Mingkun, Pavlick Ellie and Callison-Burch Chris (2014) Are Two Heads Better than One? Crowdsourced Translation via a Two-Step Collaboration of Non-Professional Translators and Editors. *The 52nd Annual Meeting of the Association of Computational Linguistics* 2014/6. Available at: http://www.cis.upenn.edu/~ccb/publications/crowdsourced-translation-via-collaboration-between-translators-and-editors.pdf.
Keywords: open and collaborative translation.

Yong, Lim Tek (2009) Collaborative awareness for translation groupware. In *International Conference on Information and Multimedia Technology, 2009. ICIMT'09,* 47-51.
Keywords: open and collaborative translation.

Zaidan, O. F. (2009) Z-MERT: A Fully Configurable Open Source Tool for Minimum Error Rate Training of Machine Translation Systems. *The Prague Bulletin of Mathematical Linguistics* 91, 79-88.
Keywords: open tools.

Zetzsche, J. (2011) Individual translators and data exchange standards. TAUS International Writers Group. Available at: http://www.translationautomation.com/perspectives/individualtranslatorsanddataexchangestandards.html.
Keywords: open standards and formats.

Notes on Contributors

Marco Agnetta

Marco Agnetta M.A. is research and teaching associate at the chair for Applied Linguistics, Translation and Interpreting Studies (Romance languages) at the Saarland University, Germany. His research focuses on poly semioticity in translation and translation theory (by analyzing the complex structure of opera, sitcoms etc.) as well as on the Open Access principle in linguistic and translatological research practice. Marco Agnetta acts as coordinator of the re-search center Hermeneutics and Creativity (www.hermeneutik-und-kreativitaet.de) established on the above mentioned chair in 2012.

Amparo Alcina

Amparo Alcina is a Senior Lecturer at the Universitat Jaume I of Castellón (Spain) where she teaches Translation Technology and Terminology to translators. Her research work is about translation, terminology and language technologies.

Erik Angelone

Erik Angelone is Associate Professor of Translation Studies at Kent State University's Institute for Applied Linguistics. He received his Ph.D. in Translation Studies from the University of Heidelberg. His current research interests include process-oriented translator training, cognitive processes in translation, and translation assessment. He recently co-edited a volume titled Translation and Cognition (John Benjamins 2010) with Dr. Gregory Shreve.

Marta García González

Degree and PhD in Translation and Interpreting by the University of Vigo. MA in Foreign Trade by the University of Vigo. She was a professional translator from 1997 to 2010, specializing in legal and business Translation. Since 2001, she has been a lecturer of legal and business translation at the Faculty of Philology and Translation of the University of Vigo, where she was the Director of the faculty's MA in Multimedia Translation of that university from 2010 and 2012. Currently, she is a member of the Academic Commission of the Master in Multimedia Translation and of the PhD Program in Communica-

tion. She is a member of the GETLT research group and her main research interests are legal and business translation, translation pedagogy, translation from and into minorized languages, and screen translation.

GETLT

The group GETLT (Grupo de Estudos das Tecnoloxias Libres da Tradución) (http://webs.uvigo.es/getlt) was created in 2007 at the University of Vigo, Spain, with the following goals in mind: to analyze and promote the use of free software in professional translation practice, as well as in translator training; to promote the visibility of the work done by volunteer translators of free projects; and to encourage cooperation by students, teachers and trans-lation professionals with communities involved in translating free software projects. The group is involved in the coordination of the Master in Screen Translation (http://webs.uvigo.es/multitrad) and of the PhD Program in Communication (http://webs.uvigo.es/comunitrad) of the University of Vigo.

Silvia Flórez Giraldo

Silvia Flórez Giraldo is an English-French-Spanish translator graduated from Universidad de Antioquia, Colombia. She completed an M.Phil. in English (linguistics) at the University of Bergen (Norway), and a Master in Translation and Localization Technologies at Universitat Jaume I (Spain). She also holds a Ph.D. in translation and knowledge society from Universitat Jaume I, where she wrote a dissertation on free/open-source translation technologies and their evaluation. She currently works as an independent translator and teaches at her home university.

Cristian Lakó

Cristian Lakó is a junior researcher and member of the academic staff at "Petru Maior" University, Tg. Mureș, Romania. He received his PhD from "Al. I. Cuza" University, Iași, Romania, in 2014. His current research interests revolve around website translation and localization, on-line marketing and its applicability to translation studies. He is also a big IT enthusiast and enjoys writing code for his personal projects.

Adrià Martín-Mor

Adrià Martín-Mor is a lecturer in translation technologies at the Department of Translation, Interpreting and East Asian Studies at the Universitat Autònoma de Barcelona, where he teaches translation technologies and coordinates the Tradumàtica MA course. His research interests are CAT tools, machine translation and FOSS software.

Philipp Neubauer

Philipp Neubauer currently works as freelance translator, interpreter (EN-DE), terminologist and language instructor (EFL). He is also an independent researcher in the fields of translation studies and terminology research and offers consulting services to language industry professionals and enterprises. After studying and completing the Bavarian state exam for translating and interpreting at the Munich Municipal College of Translating, Interpreting and Foreign Languages (FIM), he obtained further degrees from the University of Salford (UK): a distinction-level Master's and a PhD in modern languages. His research interests include the philosophy of information, especially as applied to practical problems, knowledge management and the sociology of the emergent multilingual knowledge society.

Ramon Piqué i Huerta

Dr. Ramon Piqué i Huerta has been a lecturer on IT and translation technologies at the Department of Translation, Interpreting and East Asian Studies at the Universitat Autònoma de Barcelona since 1993. His fields of research are the analysis of the digitalized translation process and the tools used. He is a member of the Tradumàtica research group and the director of the research e-journal Revista Tradumàtica on translation technologies.

Pilar Sánchez-Gijón

Dr. Pilar Sánchez-Gijón is a senior lecturer in translation technologies at the Department of Translation, Interpreting and East Asian Studies at the Universitat Autònoma de Barcelona, where she teaches subjects related to CAT tools, corpus linguistics, machine translation, post-editing, and terminology.

Peter Sandrini

Peter Sandrini is currently attached to the Department of Translation Studies at the University of Innsbruck as an Assistant Professor. He has published on legal terminology, translation, website localization and translation technology and holds courses on the same subject fields. He is also the initiator of USBTrans and *tuxtrans*, a project which aims at bringing Open Source software to translators training and the translator's desktop.

Tradumàtica Group

The Tradumàtica research group (www.tradumatica.net) gathers researchers interested in translation technologies. The group coordinates the Tradumàtica MA course (www.master.tradumatica.net), publishes the e-journal Revista Tradumàtica (www.revista.tradumatica.net) and carries out research projects (http://grupsderecerca.uab.cat/tradumatica/en/content/projects), such as the TRACE project, on the effect of CAT tools, and ProjecTA (www.projecta. tradumatica.net), on Statistical Machine Translation.

María Teresa Veiga Díaz

Degree and PhD in Translation and Interpreting by the University of Vigo. She worked as a professional translator from 1997 to 2011, mainly in the field of scientific translation. Since 2003 she has been a lecturer of scientific translation at the Faculty of Philology and Translation of the University of Vigo. She was the Director of the MA in Multimedia Translation of the University of Vigo until 2015 and still part of its Academic Commision. She is a member of the GETLT research group and her research interests include scientific translation, translation pedagogy, and multimedia translation and minorized languages.

Index

Thinking about openness and implementing openness in our attitudes and actions have considerable bearing on our conception of ourselves as translators or researchers. Openness indeed questions the very role of translated texts, multilingual translation resources, the ethics of translators, their professional behavior, the self-conception of academics and researchers, as well as the role and availability of research results in society. It therefore constitutes one of the most stimulating challenges that the world of professional translation and translation studies have yet faced.

The contributions to this volume review some of these topics in three thematic sections: the first and most substantial part deals with the concept of openness in ICT (open data, open tools, open computer systems, and quality evaluation of open software), the middle part is concerned with translators training and the use of open software, and the last part discusses openness in academia on the basis of the concepts of a Scientist 2.0 and Digital Scholarship. An exhaustive list of references covering the topic is given as an appendix, as well as a keyword index.