Taking Statistical Machine Translation to the Student Translator

Stephen Doherty
Centre for Next Generation Localisation,
Centre for Translation and Textual Studies,
Dublin City University
stephen.doherty@dcu.ie

Dorothy Kenny
Centre for Next Generation Localisation,
Centre for Translation and Textual Studies,
Dublin City University
dorothy.kenny@dcu.ie

Andy Way
Applied Language Solutions
Delph,
Oldham,
OL3 5FZ,
United Kingdom
andy.way@appliedlanguage.com

Abstract

Despite the growth of statistical machine translation (SMT) research and development in recent years, it remains somewhat out of reach for the translation community where programming expertise and knowledge of statistics tend not to be commonplace. While the concept of SMT is relatively straightforward, its implementation in functioning systems remains difficult for most, regardless of expertise. More recently, however, developments such as SmartMATE have emerged which aim to assist users in creating their own customized SMT systems and thus reduce the learning curve associated with SMT. In addition to commercial uses, translator training stands to benefit from such increased levels of inclusion and access to state-of-the-art approaches to MT. In this paper we draw on experience in developing and evaluating a new syllabus in SMT for a cohort of post-graduate student translators: we identify several issues encountered in the introduction of student translators to SMT, and report on data derived from repeated measures questionnaires that aim to capture data on students' self-efficacy in the use of SMT. Overall, results show that participants report significant increases in their levels of confidence and knowledge of MT in general, and of SMT in particular. Additional benefits – such as increased technical competence and confidence – and future refinements are also discussed.

1 Introduction

Statistical Machine Translation (SMT) (e.g. Koehn, 2010) is based on an intuitively simple strategy: rather than work out how to translate from one language to another, try to learn from what human translators have already done (Hearne and Way, 2011). However, despite the simplicity of the idea and the fact that SMT actually uses human translations as data, SMT quickly becomes difficult for translators to understand given the complexity of the statistical models it uses in training, and the nature of the algorithms it uses to generate the most likely translation at runtime. This is disempowering for human translators, as they are not generally in a position to contribute to the development of such systems, or to their introduction in translation workflows. [1]

[1]See the following quote from Dion Wiggins, CEO of Asia Online, at http://www.linkedin.com/groups/Looks-like-licencebased-model-MT-148593.S.74453505?qid=579815d2-fdfd-46bb-ac04-3530d8808772andtrk=group_search_item_list-0-b-ttl:
"*The translator should not have any ownership in the translation process*. They are 1 part of the translation process. There is much more than the translator. Ownership should be with the level of the LSP and the client, not at the translator level. Management is required, translators perform a step - translation (sometimes well, sometimes poorly), then there is post editing, proofing, project management, quality control and much more. *Allowing the translator (who is usually a freelancer and also works for your competitors) to control the translation process is a recipe for disaster*." (Our emphasis). Note the contribution from Mirko Plitt (Autodesk) in the same discussion with respect to self-serve MT platforms such as SmartMATE: "How ironic would it be if MT of all things would help translators regain ownership of the translation process!"

Furthermore, they can find themselves confined to reactive 'after-the-event' roles in SMT (e.g. post-editing), and excluded from other, proactive, holistic roles that many translators commonly adopt in their professional lives. Such scenarios are uncomfortable for those who educate translators; while we want our students to be well-versed in the use of contemporary technologies, we quite clearly do not want them to be forced into constricted, disempowering roles, a danger proposed, by Bowker (2005), for instance, where translators run the risk of blindly following translation memory (TM) output over their judgment and experience.

At the same time, we are convinced that many translators stand to gain considerably from the use of SMT, while developers of SMT systems also stand to benefit from a greater uptake of the technology by translators, and the possibilities of gaining insights from user experiences which may help them improve their engines (Volk and Harder, 2007; Way and Hearne, 2011). Like many sociologists of technology (Pinch, 2008), we take the view that markets for technologies are actively constructed, and we acknowledge the role that both the vendors of technologies and educators play in such market construction.

Translation technology has grown out of the domain of translation studies and evolved into a specialization in its own right (Snell-Hornby, 2006; Alcina, 2008), but is largely neglected within the mainstream theories of translation (Munday, 2009). One of the integral – but currently missing – pieces is a published syllabus that educators can use to teach translation students about MT in a way that empowers rather than instrumentalizes them in MT workflows, especially SMT workflows. Such a syllabus would include theoretical components tailored to meet the needs of students who are not majoring in computer science. The theoretical content would underpin the practical components in which students learn: how to gather appropriate data, how to train an SMT system; how to improve system performance; how to evaluate SMT output; etc. Many existing discussions on the teaching of MT either pre-date the rise of SMT (e.g. Kenny and Way, 2001; Somers, 2003) or embed the discussion of MT in the wider context of teaching translation technology (see O'Brien and Kenny, 2006) and are thus necessarily limited in their focus on SMT. To our knowledge, no other sources have attempted to systematically *evaluate* effective learning in a module on SMT.

In this paper we present the first results from the evaluation of a combined teaching and research project which aims to produce such a syllabus.[2] In the project, conducted in the first half of 2012, thirty eight students taking Masters-level translation programmes at Dublin City University (DCU) took a course of lectures and practical sessions (or *labs)* on: MT history and early development, the concepts behind MT with an emphasis on SMT; MT evaluation using human and automatic metrics; the use of SMT in translation workflows; and the roles of humans in SMT workflows. In labs and in their take-home assignment, students used the self-serve SMT package SmartMATE (Way *et al*., 2011) [3] developed and hosted by Applied Language Solutions, to create and optimize their own SMT systems. SmartMATE was considered ideal for use in this experiment, as it did not require students to have the kind of programming knowledge required to install freely available open-source solutions such as Moses (Koehn *et al*., 2007), but it did allow them considerable freedom to build and customize their own SMT systems.

The evaluation of the SMT syllabus at DCU was based on a number of different instruments including participant questionnaires, student and lecturer logs, end-of-module assignments, and focus groups. In this paper we report principally on data derived from the questionnaires. The questionnaires were designed to elicit the usual demographic information about participants (age, sex, etc.), as well as information about their normal levels of computer use and their experience using MT in particular.[4] Crucially, the questionnaire also aimed to capture information about the students' *self-efficacy* in the use of SMT. In the rest of this paper we first describe the self-efficacy construct (Section 2). We then go on to describe SmartMATE, the system used for practical work in labs and the take-home assignment in Section 3. The methodology adopted in the evaluation is set

[2] The details of the refined syllabus, post-evaluation, will be published in due course.

[3] http://www.smartmate.co

[4] All students involved were already familiar, for example, with translation memory, having already covered this technology in their programme.

out in detail in Section 4. Results and discussion follow in Sections 5 and 6, respectively. Our conclusions and ideas for further refinements are set out in Section 7.

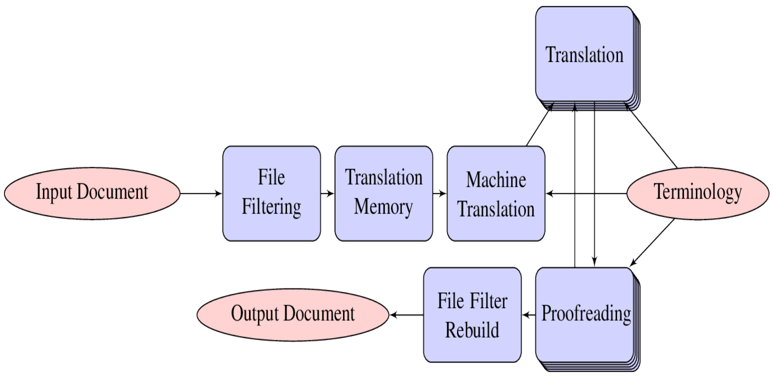

Figure 1: Typical translation workflow in SmartMATE

2 Self-Efficacy

Self-efficacy is the principal component of social cognitive learning theory, a paradigm that posits models as the principal source for learning new behaviours and instigating behavioural change (Bandura, 1977; Simz and Manz, 1982). The construct of self-efficacy pertains to an individual's – in this case the learner's – confidence in their ability to control their own thoughts, feelings and actions to produce a desired outcome (Bandura, 1986). The predominant *modus operandi* of most social cognitive processes is observational learning, which occurs via:

a) Demonstration of the desired behaviour to the learner (directly and indirectly);
b) Instruction of described behaviour to the learner by verbal means;
c) Symbolic - the use of other media and multi-modal approaches to convey the behaviour.

In the context of our teaching scenario, students learn by a combination of all three: (a) direct demonstration of the software by the lecturer using a data projector in a well-equipped computer lab, and indirect observation of peers on the part of the student; (b) instruction in the form of verbal cues during the lectures and practical sessions - both from the lecturer and the peer group; and (c) by means of a variety of other media made available to all students via an e-learning platform, such as movies, presentations, non-academic material (e.g. commercial insights into contemporary translation technology issues), and social collaborative interaction via an online forum.

In terms of realized effects, perception of self-efficacy can influence greatly an individual's performance (Locke *et al*., 1984), decision making (Betz and Hackett, 1981), and attrition rate (Brown and Inouye, 1978). In the case of self-efficacy in the use of computer software, both direct experiences (e.g. using the software oneself), and indirect experiences (e.g. observing a peer using a software program) allow an individual to gain insight into their own ability to carry out a given set of tasks. Such insight is typically captured by means of self-report measurements. Vispoel and Chen (1990) recommend that due to the individualized nature of self-efficacy, specific measures should be used for appropriate situations rather than a more generalized approach. Compeau and Higgins (1995) provide a validated instrument for assessing computer self-efficacy in the form of a ten-item questionnaire. Other examples of self-efficacy measures related to the use of computer technologies can be found in, for example, Burkhardt and Brass (1990), Delcourt and Kinzie (1993), and Miltiadou and Yu (2000). Much support exists for the use of self-efficacy as a reliable predictor of academic performance and technical competence (Ames, 1984; Multon *et al*., 1991; Nicholls and Miller, 1994). In the context of our evaluation, we can conclude that if we observe an increase in self-efficacy after students have completed our SMT course, then this can be reliably interpreted as an indication that their technical and academic performance using SMT has indeed improved.

3 SmartMATE

SmartMATE (Way *et al*., 2011) is an online self-serve translation platform, designed as a one-stop portal where users can upload their TM files (in TMX format, which can be exported from any Translation Management System software),[5] and

[5]Translation Memory eXchange: http://www.gala-global.org/oscarStandards/tmx/tmx14b.html

create user-customized SMT engines trained using these TMs.

All of these capabilities are integrated in a typical translation workflow, as shown in Figure 1 for SmartMATE. Assume we have an input document which needs to be translated. Since there is a variety of file formats in which this document can be encoded, it is first sent to File Filtering, which produces an XLIFF[6] (XML Localisation Interchange File Format) file containing only the translatable text, without additional elements such as images or formatting information.

All of the components in SmartMATE take an XLIFF file as input and produce a modified one as output, except for File Filtering, which can accept a wide range of document formats, including Microsoft Office Suite file formats (e.g. Word, Excel and PowerPoint), as well as other popular formats such as .rtf, .html, .ttx and .txt. The XLIFF file – either originally in .xlf format, or generated from some other format via File Filtering – can then optionally be sent through the TM component in order to leverage any previous translations, and through MT for segments which do not match any TM entry at the required threshold level. At this stage, the document becomes available for post-editing. SmartMATE provides an online multi-user Editor Suite (cf. Penkale and Way, 2012). Users can make use of the editor themselves to translate the document, or they might delegate this to a third party. After translation has finished, the translated XLIFF file is sent back to File Filtering in order to recover the original file format.

TMX files can be used in two different ways in SmartMATE. Firstly, they can be used as traditional Translation Memories. When a new document is ready to be translated, segments in the document which exactly match any TM entry will appear in the editor suite as pre-translated using the target side of that entry. In addition to exact matches, SmartMATE allows fuzzy and in-context exact matches. After the document has been translated and accepted by the proofreader, TMs may be automatically updated to include the newly translated content.

Secondly, a user's TMX files (and glossaries, if available) can be used to train an SMT engine completely automatically. Plain bilingual text is extracted from the TMX files to create a parallel corpus, which is then subjected to multiple stages of corpus clean-up, one of the main reasons why SmartMATE manages to considerably outperform the Moses (Koehn *et al*., 2007) baseline on which it is built. Once the various models (phrase-based translation model, language model, lexicalized reordering model) have been constructed and tuned via MERT (Och, 2003), the engine is ready to be tested.

A recent example of how effective SmartMATE can be was where a new user (availing of the 30-day free trial) uploaded a 1 million-segment TM file, had an English-to-Spanish engine built in less than 5 hours, and then translated 180,000 words in less than 2 hours at a rate of over 1500 words/minute, with a BLEU score (Papineni *et al*., 2002) of over 70 on a 1000-sentence held-out test set. By any measures, this is impressive, and demonstrates how effective SmartMATE can be for LSPs and individual translators alike.

4 Methodology

A mixed-methods approach was taken to the research to access rich qualitative data about the subjective experiences of the student translators, and to measure student learning using standard quantitative instruments. As already indicated, data were collected using participant questionnaires (containing items accessing experience of other translation technology tools, computer usage, and self-efficacy (see Section 5 below) in the use of SMT), translator and lecturer logs, end-of-module assignments, and focus groups. Principally, we report here on data derived from the questionnaires. These were distributed as hard-copies at the beginning of the SMT course (the first time point, *t1*) and upon its completion (the second time point, *t2*); both versions were identical in their content and appearance.

Self-efficacy was measured using a ten-item questionnaire (adopted from Compeau and Higgins, 1995), where reliability was found to be very acceptable (α = .83). Participants were given ten examples of contextually relevant situations where they were asked to rate their ability to complete a task using an SMT system. Example situations included the following: "I could use an SMT system, if someone showed me how to do it first", or "I could use an SMT system, if I had only

[6] https://www.oasis-open.org/committees/xliff/

the software manuals for reference”. Ratings for each example were on a range from 1 to 10, across three bands: not at all confident (1 – 3), moderately confident (4 – 6), and totally confident (7 – 10). Furthermore, each response was classified by the participants as a positive or negative statement. To use the above examples once again, one could state that one felt moderately confident in using SMT effectively if someone guided the way first, but moderately confident that one could *not* use the SMT package using only the Help provided as a reference.

The material covered in the course of the module can be summarized into the following areas:

I. Introduction to MT systems: a historical overview of MT, description of prominent systems, and familiarity with contemporary research and commercial applications;
II. Introduction to MT evaluation techniques: reverse engineering, and test suites;
III. Using SMT systems: identification of corpora and data types, quality assessment, training the system, customization;
IV. Automatic evaluation metrics: understanding and using a variety of metrics,[7] and combining this approach with human evaluation;
V. Error typologies for human evaluations: existing and customized models, e.g. the model proposed by Vilar *et al.* (2006);
VI. Pre- and post-processing: using and creating style guides, controlled language, theory and practice of post-editing;
VII. MT workflows: using above elements together and following a translation from start to finish as demonstrated via the take-home assignment.

For the take-home assignment, students were given the option of making use of the SMT engines that they had already created in labs during the taught module, or creating new engines. Regardless of which engine was used, students were required to furnish a description of their training corpus and its quality, and of the steps taken to train their system. Using their own engine, students translated two texts of their choosing, one in-domain and one out-of-domain (approximately 500 words each), where ‘in-domain’ and ‘out-of-domain’ are relative to the corpora used for training the systems in question. (Most students used texts from the legal or IT domains to train their systems.) They then conducted a human evaluation using an error typology and/or measures of adequacy and fluency, and an automatic evaluation using two of the aforementioned automatic evaluation metrics. Having identified errors and areas requiring improvement using these criteria, they then constructed a set of preprocessing guidelines, mostly in the form of controlled language rules. These guidelines were used to edit the two source texts, which were then re-translated using the same engine as before. These two new ‘controlled’ translations were then evaluated using the same means as before, and students commented on how, and why, the quality of the MT improved/diminished following their interventions. Finally, students provided a critical review of their methodology and recommendations for their future use of SMT engines in MT workflows such as that in Figure 1.

5 Results

Twenty nine students took part in the survey at both the beginning and end of the SMT course (females = 21, males = 8, average age = 26 years). Several language pairs were represented in the cohort, and some participants had two or three language pairs (see Table 1).

Language Pair	*n*	Language Pair	*n*
FR—EN	16	EN—SV	1
DE—EN	5	EN—IT	1
ES—EN	4	DE—FR	1
EN—ES	4	EN—CA	1
GA—EN	3	FR—PT	1
JP—EN	3	EN—LT	1
EN—DE	2	FR—LT	1
EN—PT	2	**Total**	**46**

Table 1: Language Pairs[8] by Number in Sample

[7] LED, TER, WER, PER, BLEU, NIST, METEOR, ROUGE, GTM.

[8] GA = Irish, CA = Catalan, LT = Lithuanian.

We now present results from related factors of computer use, knowledge of translation memory, and knowledge of machine translation, as we believe they are greatly influential on the students' self-efficacy specific to SMT, as described in points III and IV below.

I. Computer Use

- *Item – "How much time, on average, do you spend using a computer per week?"*
- *Scale – 1 = 0 hours, 2 = 1 to 10, 3 = 11 to 19, 4 = 20 to 39, 5 = 30+.*

On the 5-point Likert scale administered at the beginning of the SMT module, participants reported spending between "11 to 19 hours" using a computer on an average week (*t1* median = 4, mean = 4.20, SD = .71). This value decreased very slightly at the end of the module (*t2* median = 4, mean = 4.15, SD = .88). Unsurprisingly, a repeated measures t-test found no significant change between the two time points, where t = 4.38, df = 19, p = .666. In other words, the amount of time the students report spent using a computer did not change as captured at the two intervals.

II. Translation Memory

- *Item – "How would you rate your knowledge of translation memories (TMs)?"*
- *Scale – 1 = Poor, 2 = Below Average, 3 = Average, 4 = Above Average, 5 = Excellent*

For the second item where participants were asked to rate their knowledge of TMs, they report that they had an "average" level of knowledge (*t1* median = 3, mean = 3.20, SD = .81). This increased, but not significantly so, at the end of the module (t = -1.831, df = 19, p = .0828), to the "above average" level (*t2* median = 4, mean = 3.50, SD = .69).

III. General MT Knowledge

- *Item – "How would you rate your knowledge of machine translation (MT)?"*
- *Scale – 1 = Poor, 2 = Below Average, 3 = Average, 4 = Above Average, 5 = Excellent*

Reported knowledge of MT in general showed a significant increase (t = -3.322, df = 19, p = .004), where *t1* (median = 3, mean = 2.80, SD = .616) resulted in an "above average" level, and *t2* (median = 4, mean = 3.45, SD = .686) showed an improvement to "excellent".

IV. Experience

- *Item – "Do you have any professional experience with machine translation?"*
- *Scale – Yes / No*

With regard to professional experience with MT, two participants indicated positively. When asked to give details, such experience was related to very short-term and recent university projects and was therefore deemed not to affect the responses given by these participants.

V. Self-Efficacy for SMT

A repeated measures t-test found a significant difference (t = -2.276, df = 21, p = .03), where self-efficacy levels increased from an average of 59.18 (SD = 24.453) to 70.09 (SD = 16.133), where both values are scored out a maximum value of 100 – see Figure 2. Additionally, while the *t1* scores reported 16 negative situations – where participants were confident to whatever extent that they could *not* perform the given task – this number fell to just 4 for *t2*.

It should be noted that eight of the participants indicated that although they were asked to give responses for SMT in general, they had already used freely available online systems such as Google Translate[9] and/or Systran[10], and felt that they could only relate their responses directly to their experiences of the latter. Reductions in self-efficacy scores were found for all participants in this case, and although the decrease in scores was found to be insignificant (t = .979, df = 7, p = .36), where the mean of *t1*, 78.88 (SD = 18.635), dropped to 68.00 at *t2* (SD = 23.458), it is nevertheless an important consideration which will be returned to in our discussion. These participants were the only cases where decreases of self-efficacy scores were found. Qualitative explanations from the participants indicate that

[9] http://translate.google.com
[10] http://www.systran.co.uk

their initial exposure to and understanding of other SMT systems was via a more simplified interface, see for example Figure 3. In this context, the user is limited in their interaction with the system. However, upon sourcing training data, building their own systems, and navigating though the syllabus, participants gained greater insight into the amount of work and level of expertise inside the '*black box*'.

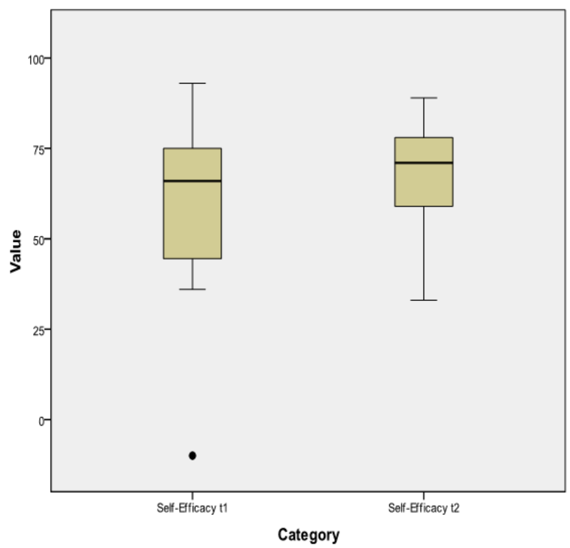

Figure 2: Self-Efficacy at Both Time Intervals

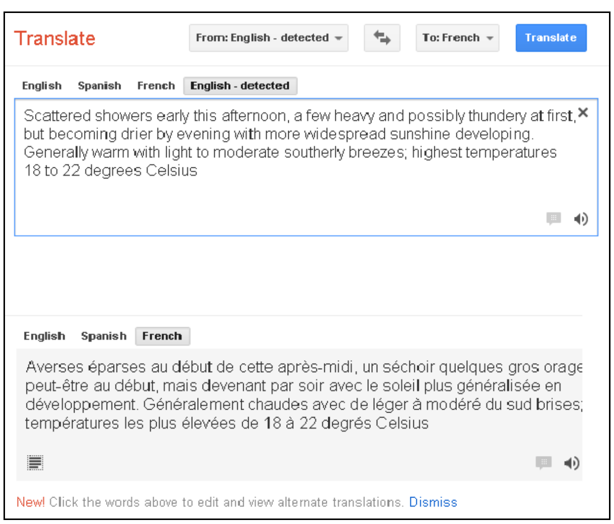

Figure 3: Google Translate Window (translate.google.com)[11]

Furthermore, using a hierarchical multiple regression analysis, both computer use (β = .451, p = .025) and knowledge of MT (β = .446, p = .026) were identified as significant predictors of SMT self-efficacy. Together, both variables explained 55.1% of the variance (r^2 = .601, F = (16, 12) 12.059, p = .001). In other words, using a computer regularly and being knowledgeable of MT in general, leads to greater self-efficacy for SMT specifically.

Lastly, as evident from Table 2 and Figure 4, we find significant correlations between knowledge of TMs and MT, and of both scales with self-efficacy. While computer use does not correlate significantly with knowledge of TMs or MT, it has a moderate positive correlation with self-efficacy. These findings add support to the measurement of self-efficacy, and each scale's construct validity, e.g. if knowledge of TM did not correlate with MT or self-efficacy, it would question how the concept is measured and/or is defined.

Computer Use	-	.342	.383	.677**
Knowledge of TM	.342	-	.926**	.572*
Knowledge of MT	.383	.926**	-	.603**
Self-Efficacy	.677**	.572*	.603**	-

Table 2: Correlations (ρ) for Scale Variables[12]

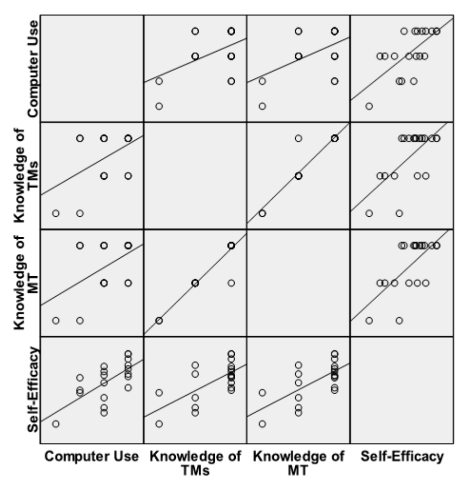

Figure 4: Correlation of Scale Variables (ρ)

[11] For spacing reasons, the source and target windows are displayed here vertically; their actual presentation is side-by-side.

[12] ** Correlation is significant at the 0.05 level (2-tailed); * correlation is significant at the 0.01 level (2-tailed).

6 Discussion

While the findings support our initial hypothesis that there would be increased levels of self-reported self-efficacy in relation to SMT, several interesting topics arise throughout the process of this study. The issue of construct validity is brought to the fore where at *t1* some students had used online SMT systems such as Google Translate and replied to the questionnaire based only on their experiences with these systems. It must be noted that while online systems tend to be very user-friendly, they allow very limited access to the *engine* itself, but rather offer a simple window for inputting text, and by means of a button click, accessing its candidate translation(s). In retrospect, the questionnaire could have reflected this and named SMT systems specifically; however, this may not be a valid solution given that participants may generalize from whatever SMT exposure they have experienced from one system and relate that to another. Therefore, it appears logical to simply remove such data from the analyses as employed here. However, it is also necessary to ask participants more explicitly if they have experience with online SMT systems, and to report on those separately.

7 Conclusion

From the above findings it is evident that there were significant improvements in reported efficacy in relation to SMT specifically, and while computer usage did not change, improvements were found for general MT knowledge and (to a lesser extent) TMs. In light of these results we can demonstrate student and translator empowerment, where the role of the translator vis-à-vis SMT is made more active rather than passive. As improvements in self-efficacy are not specific in their benefits (e.g. they spill over to other areas such as general IT skills and technical competence (Webster *et al*., 1990)), such an experience is especially relevant to anxious computer users – in this context translators presented with many technical challenges – as it ensures greater familiarity with addressing problems, and an understanding from the developers' perspective.

Future work will continue our exploration of the use of the online self-serve MT paradigm in the teaching of MT and related translation technologies. Of importance is that like SmartMATE, any system we use is welcoming to non-programmers and allows students access to a user-friendly environment where they can put into practice the knowledge they have acquired from their studies in translation and MT. Findings from the study have been used in the development of a formal curriculum for the coming academic year, which will be evaluated once again, and its findings will be fed back to the translation and MT communities. We intend to provide more reliable means of using related MT tools, such as automatic evaluation metrics, for students of MT and translators in general. Additionally, it is hoped that by making MT more accessible to students of translation, they can become more active in translation technology, and perhaps add to MT and related research in future; we also hope to have provided students with many cutting edge skills for their own employment opportunities in the translation, localisation, and related industries.

What is also of interest is the diverse nature of cohorts of student translators who are typically proficient in two or more languages and, in our case, have relocated to a country speaking one of their source languages. Intercultural differences are likely to be evident in such cases, especially in more sensitive psychological measures such as self-efficacy and student-teacher interactions (Oettigen, 1995; Scholz *et al.,* 2002).

Lastly, close cooperation between DCU and Applied Language Solutions also meant that students could be integrated into a feedback loop, receiving useful explanations and insights from the SmartMATE team. Having so many users access the tool suite synchronously provided useful load testing for ALS, and beneficial intensive feedback on various aspects of the tool in field testing by plausible end-users.

Acknowledgments

This research is supported by the Science Foundation Ireland (Grant 07/CE/I1142) as part of the Centre for Next Generation Localisation (www.cngl.ie) at Dublin City University. The authors would like to thank Dr. Jie Jiang of Applied Language Solutions for his support throughout this project, and the students who participated.

References

Amparo Alcina. 2008. Translation technologies: scope, tools and resources. *Target*, **20**(1): 79–102.

Albert Bandura. 1977. *Social Learning Theory*. Englewood Cliffs, NJ: Prentice-Hall.

Albert Bandura. 1986. *Social foundations of thought and action: A social cognitive theory*. Englewood Cliffs, NJ: Prentice-Hall.

Nancy E. Betz and Gail Hackett. 1981. The relationships of career-related self-efficacy expectations to perceived career options in college women and men. *Journal of Counseling Psychology*, **28**: 399-410.

Lynne Bowker. 2005. Productivity vs Quality? A pilot study on the impact of translation memory systems. *Localisation Focus*, **4**(1): 13–20.

Irvin Brown and Dillon K. Inouye. 1978. Learned helplessness through modeling: The role of perceived similarity in competence. *Journal of Personality and Social Psychology*, **36**: 900-908.

Marlene E. Burkhardt and Daniel J. Brass. 1990. Changing patterns or patterns of change: The effects of a change in technology on social network structure and power. *Administrative Science Quarterly*, **35**: 104-127.

Ames, Carole. 1984. Achievement attributions and self-instructions under competitive and individualistic goal structures. *Journal of Educational Psychology*, **76**: 478-487.

Deborah R. Compeau and Christopher A. Higgins. 1995. Computer Self-Efficacy: Development of a Measure and Initial Test. *MIS Quarterly*, **19**(2): 189-211.

Marcia A. Delcourt and Mable B. Kinzie. 1993. Computer technologies in teacher education: The measurement of attitudes and self-efficacy. *Journal or Research and Development in Education*, **27**: 35-41.

Mary Hearne and Andy Way. 2011. Statistical Machine Translation: A Guide for Linguists and Translators. *Language and Linguistics Compass,* **5**: 205-226.

Dorothy Kenny and Andy Way. 2001. Teaching Machine Translation and Translation Technology: a Contrastive Study. In *Proceedings of the MT Summit VIII Workshop on Teaching Translation, Santiago de Compostela, Spain,*13-17.

Philipp Koehn. 2010. *Statistical machine translation*. Cambridge: Cambridge University Press.

Philipp Koehn, Hieu Hoang, Alexandra Birch, Chris Callison-Burch, Marcello Federico, Nicola Bertoldi, Brooke Cowan, Wade Shen, Christine Moran, Richard Zens, Chris Dyer, Ondřej Bojar, Alexandra Constantin and Evan Herbst. 2007. Moses: open source toolkit for statistical machine translation. In *ACL 2007: proceedings of demo and poster sessions*, Prague, Czech Republic, pp.177--180.

Kristiina Kumpulainen and David Wray. 2002. *Classroom Interaction and Social Learning: From Theory to Practice*. New York, NY: Routledge Falmer.

Edwin A. Locke, Elizabeth Frederick, Cynthia Lee and Philip Bobko. 1984. Effect of self-efficacy, goals, and task strategies on task performance. *Journal of Applied Psychology*, **69**: 241-251.

Marios Miltiadou and Chong Ho Yu. 2000. Validation of the online technologies self-efficacy scale (OTSES). *Proceedings of SITE 2000, Society for Information Technology and Teacher Education International Conference*, February, San Diego, CA: 1-17.

Karen D. Multon, Steven D. Brown and Robert W. Lent. 1991. Relation of self-efficacy beliefs to academic outcomes: A meta-analytic investigation. *Journal of Counseling Psychology*, **38**: 30-38.

Jeremy Munday. 2009. *The Routledge Companion to Translation Studies.* London: Routledge.

Joe D. Nicholls and Raymond B. Miller. 1994. Cooperative learning and student motivation. *Contemporary Educational Psychology*, **19**: 167-178.

Sharon O'Brien and Dorothy Kenny. 2006 Translation technology in the translator training syllabus. In: Anne Gallagher and Muiris Ó'Laoire (Eds.), *Language Education in Ireland: Current Practice and Future Needs* (pp. 185-192). Dublin: Irish Association for Applied Linguistics.

Franz Josef Och. 2003. Minimum error rate training in statistical machine translation. In *41st Annual Meeting of the Association for Computational Linguistics, Proceedings of the Conference*, Sapporo, Japan, pp. 160--167,

Gabriele Oettigen. 1995. Cross-cultural perspectives on self-efficacy. In: Albert Bandura (Ed.), *Self-efficacy in changing societies* (pp. 149-176). New York: Cambridge University Press.

Kishore Papineni, Salim Roukos, Todd Ward and Wei-Jing Zhu. 2002. Bleu: a method for automatic evaluation of machine translation. In *40th Annual Meeting of the Association for Computational Linguistics*, pages 311–318, Philadelphia, PA.

Sergio Penkale and Andy Way. 2012. An Online End-To-End MT Postediting Framework. In

Proceedings of AMTA Workshop on Post-editing, San Diego, CA. (under review).

Trevor Pinch. 2008. Technology and Institutions: living in a material world. *Theory and Society*, **37**: 461–48.

Urte Scholz, Benicio G. Doña, Shonali Sud and Ralf Schwarzer. 2002. Is general self-efficacy a universal construct? Psychometric findings from 25 countries. *European Journal of Psychological Assessment*, **18**(3): 242-251.

Henry P. Sims and Charles C. Manz. 1982. Social Learning Theory. *Journal of Organizational Behavior Management*, **3**(4): 55-63.

Mary Snell-Hornby. 2006. *The Turns of Translation Studies*. Amsterdam and Philadelphia: John Benjamins.

David Vilar, Jia Xu, Luis F. D'Haro and Hermann Ney. 2006. Error analysis of statistical machine translation output. *Proceedings of the Fifth International Conference on Language Resources and Evaluation*. Genoa, Italy 697-702.

Walter P. Vispoel and Ping Chen. 1990. Measuring self-efficacy: The state of the art – Report No. TM 017588). *American Educational Research Association*, Boston, MA (ERIC Document Reproduction Service No. ED 338 712).

Martin Volk and Søren Harder, S. 2007. Evaluating MT with Translations or Translators. What is the Difference*? Proceedings of MT Summit XI.* Copenhagen, Denmark, 499–506.

Andy Way and Mary Hearne. 2011. On the Role of Translations in State-of-the-Art Statistical Machine Translation. Language and Linguistics Compass, **5**: 227-248.

Andy Way, Kenny Holden, Lee Ball and Gavin Wheeldon. 2011. SmartMATE: Online Self-Serve Access to State-of-the-Art SMT. *Proceedings of the Third Joint EM+/CNGL Workshop "Bringing MT to the User: Research Meets Translators", JEC 2011*, Luxembourg: 43-52.

Jane Webster, J. B. Heian and Jeffrey E. Michelman. 1990. Mitigating the Effects of Computer Anxiety through Training. *Proceedings of the Eleventh International Conference on Information Systems, December 16-19, 1990, Copenhagen, Denmark:* 171-182.