



Machine Translation & Translator Training: Exploration of Students' Abilities and Needs

Khetam Al Sharou

PhD researcher in Translation Studies

UCL, Centre for Translation Studies (CentraS)

Khetam.Sharou.14@ucl.ac.uk

Translation Technologies in Translator Training

- Systems adopted by most of translation programmes are limited to certain expensive commercial packages of CAT such as SDL Trados Studio, MemoQ, and Déjà Vu.
- MT has only been included in the broader context of translation teaching technology (Kenny and O'Brien 2006).
- Proposals for incorporating MT into translation programmes are limited to those of post-editing as the one made by O'Brien (2006, 2002); O'Brien, *et al.* (2014) and Torrejón and Rico (2002).

Growing interest:

Machine Translation and Statistical Machine Translation

- Recent developments in MT, data-driven SMT systems, have caused consternation within the academic community and changed the ways in which new technologies are utilised by translation agencies (Austermuehl 2013).
- From SMT as domain of large corporations (Google, Microsoft, etc.)
- To individuals
- Open-source Moses toolkit (Koehn *et al.* 2007: 178).

What is Moses?

- ❖ A freely accessible MT engine – FP7 (Koehn 2014; Koehn et al. 2007)
- ❖ Primary development platform for Moses is Linux.
- ❖ Considerable freedom to build and customize your own SMT systems
- ❖ Can be trained for any language pair automatically

Recent Trends

- Incremental revisions of training needs
- Emerging needs for integrating SMT teaching into translators' training (Doherty *et al.* 2012; Austermuehl 2013).

The Role of Translator in the SMT Workflows

Limited - post-editors

NO

Translators and SMT

YES

- Austermuehl points out that:

‘any current discussion of how to integrate the teaching of translation technology solutions, or tools, into the training of future translators needs to acknowledge in particular the impact that recent developments in MT, especially in SMT has had, and will continue to have, on the lives of professional translators’ (2013: 328).

Syllabus SMT Proposals

SMT syllabus: technologies and human interactions

- Limitations:
 - Translators as post-editors
 - Passive receivers
- Empowering alternatives

The Proposed Syllabus

The research aims to find answers to the following questions:

1. Would it be possible to integrate the FOSS SMT into an MA translation programme?
 - a. If the only obstacle is training on Linux OS, could this be achieved with a set of simple, introductory sessions on this FOSS?
2. What tasks have to be included in the syllabus (Content) and how will the tasks be used in the classroom (Teaching Methodology)?
3. Will the trainee translators learn the required skills to run the system under consideration?
4. If yes, what will this mean and what type of learning does it entail?

The Proposed Syllabus

This research sets out to test the hypothesis that Moses FOSS SMT can be integrated into translator training programmes. The overall data analysis aims to provide answers as to the possibility of a wider integration of Moses with other translation technologies into translation programmes in Higher Education (HE).



Research Methodology (Preparation Stage)

- a) Me, testing the system and preparing the teaching material
- b) MA-level students participated in the data collection, volunteering to train on Moses MT

The Design of the Syllabus

- A task-based approach is adopted to design the tested MT training syllabus
- Using this approach, the tested syllabus was divided into eight theoretical session and hands-on lessons

Lessons Design

First lesson	An overview of the history of MT; MT approaches, linguistic issues; examples of online machine systems; statistical machine translation (details of how SMT works); the examined System.
Second lesson	Introduction to basic computing skills in Linux to ensure familiarity with its basic commands.
Third Lesson	Installation and building a Moses MT from source
Fourth Lesson	Installing essential external tools needed to the Moses setup, such as the word alignment tool Mgiza
Fifth Lesson	Sourcing and preparing data to train the MT system
Sixth Lesson	Running Moses, or the final step to create the MT
Seventh Lesson	Translating by using Moses MT
Eighth Lesson	Evaluating the quality of the MT outputs through testing the quality of their own MT system using different texts which had been previously chosen as suitable in-domain and out-of-domain test sets.

Training Module in Practice

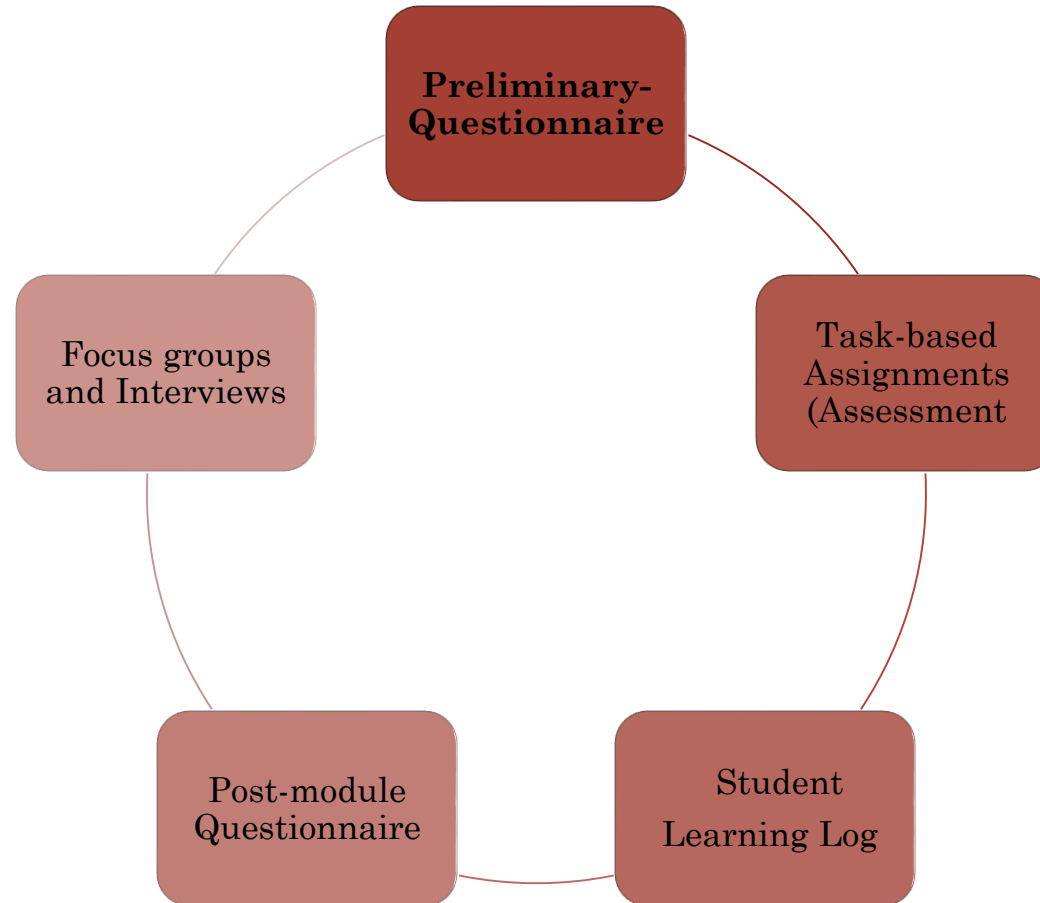
Demographic information and Language Competences

- ❖ 21 MA students
- ❖ 75% Female, 25% Male
- ❖ Arabic is their mother-tongue



Pictures used with participants' consent.

Research Methods (Methodology)



Preliminary Results

Before – After

- Participants' knowledge of MT
- Their opinions and attitude towards MT
- Self-confidence through self-efficacy measures

Self's reflections and Observations

- Tasks accomplished 62%
- A little bit complicated at the beginning (new interface, co-occurrences of errors, lack of focus)
- Mistakes were expected so they accepted them and developed their skills
- Learning became much easier and more enjoyable.

The questions are still:

What do they need?

- They need to know more MT (how it works, its limitation, etc.)
- They need to be made aware of their own ability
- What role MT technical skills can play in their future career.

What can they do?

- They can learn how to use MT, even if they have to use Linux.
- They can develop their technical skills even if they are taught to do so in a short time.

Thank you!

Some References

- Bowker, L. (2014). “Computer-Aided Translation: Translator Training”, in Chan, S. (ed.) *The Routledge Encyclopaedia of Translation Technology*, Routledge, 88-104.
- Candel-Mora, M. A. (2016). “Translator Training and the Integration of Technology in the Translator’s Workflow”, in (ed.) Carrió-Pastor, M. L., *Technology Implementation in Second Language Teaching and Translation Studies. New Tools, New Approaches*, Singapore: Springer Singapore, 49-69.
- Doherty, S., Kenny, D. and Way, A. (2012). “Taking Statistical Machine Translation to the Student Translator”, in AMTA, the Tenth Biennial Conference of the Association for Machine Translation in the Americas, San Diego, 28 Oct to 1 Nov 2012, 1-10, pdf, [online], Available at: <<http://doras.dcu.ie/17669/1/AMTA-2012-Doherty-1.pdf>> [Accessed on 01 Nov 2014].
- Kenny, D. and Doherty, S. (2014). “Statistical Machine Translation in the Translation Curriculum: Overcoming Obstacles and Empowering Translators”, *The Interpreter and Translator Trainer*, 8 (2), 276-294.
- Koehn, P., Hoang, H., Birch, A., Callison-Burch, C., Federico, M., Bertoldi, N., Cowan, B., Shen, W., Moran, C., Zens, R., Dyer, C., Bojar, O., Constantin, A., Herbst, E., (2007). “Moses: Open Source Toolkit for Statistical Machine Translation”, in Proceedings of the ACL 2007 Demo and Poster Sessions, Association for Computational Linguistics, Prague, Jun 2007, 177–180, [online], Available at: <<http://acl.ldc.upenn.edu/P/P07/P07-2045.pdf>> [Accessed on 22 Oct 2014].
- Moses Website (2015). *Welcome to Moses*, [Online], Available at; <<http://www.statmt.org/moses/>> [Accessed on 10 Sep 2014].
- O'Brien, S., Balling, L.W., Carl, M., Simard, M., and Specia, L., (eds) (2014). *Post-Editing of Machine Translation: Processes and Applications*, Cambridge Scholars Publishing.
- O'Brien, S. and Simard, M. (eds) (2014) 'Special Issue on Post-Editing'. MACHINE TRANSLATION, 28: 159-329.
- O'Brien, S. (2002). “Teaching Post-Editing: A Proposal for Course Content”, Proceedings of the 6th EAMT Workshop on “Teaching Machine Translation”, *EAMT/BCS*, UMST, Manchester, UK, 99-106, [online], Available at: <<http://mt-archive.info/EAMT-2002-OBrien.pdf>> [Accessed on 05 Apr 2014].