MACHINE TRANSLATION AND LANGUAGE LEARNING: WHAT THE RESEARCH TELLS US

Dorothy Kenny
School of Applied Language and Intercultural Studies
Research Seminar
03 February 2022
Dublin City University (online)
MultiTraiNMT

**Erasmus+ Strategic Partnership** that aims to develop, evaluate and disseminate open access materials to enhance teaching and learning about machine translation among language learners, language teachers, trainee translators, translation teachers and professional translators.

**Priorities include ensuring:**
- up-to-date syllabus in MT
- development of technical, critical and ethical skills and competences
- entrée into the world of machine learning for language and translation students, their teachers and others

**Core Partners**
- **Universitat Autònoma de Barcelona** Pilar Sánchez-Gijón & Olga Torres-Hostench
- **Universitat d’Alicant** Juan Antonio Pérez-Ortiz, Felipe Sánchez-Martínez & Mikel Forcada
- **Université Grenoble Alpes** Caroline Rossi & Alice Carré
- **Dublin City University** Dorothy Kenny, Sharon O’Brien & Joss Moorkens
- **Kantan AI** Riccardo Superbo
- **Promsigo** Gema Ramirez-Sánchez
Project Deliverables

- Open-access online book on machine translation for everyone
  Sample: [http://tinyurl.com/hownmtworks](http://tinyurl.com/hownmtworks)

- Free online activities on machine translation
  Sample: [http://tinyurl.com/nmtactivities](http://tinyurl.com/nmtactivities)

- Pedagogically-oriented open neural machine translation platform called MutNMT
  [https://ntradumatica.uab.cat/](https://ntradumatica.uab.cat/)
The Coursebook

- Module 1. Multilingualism
- Module 2. Introduction to MT
- Module 3. Evaluating machine translation quality
- Module 4. How to prepare and select texts for MT
- Module 5. Post-editing
- Module 6. Ethics and MT
- Module 7. How NMT works
- Module 8. Customisable NMT engines
- **Module 9. MT and language learning**
MT in language learning: ramifications

High-level (national/regional) language policies

Notably, in January 2012, Lawrence Summers, former president of Harvard University, commented that “English’s emergence as the global language, along with the rapid progress in machine translation make it less clear that the substantial investment necessary to speak a foreign tongue is universally worthwhile”. (Clifford et al. 2013: 109)

University policies  (e.g. Groves & Mundt 2015; Mundt & Groves 2016 )

- admission
- assessment
- academic honesty / plagiarism

Instructional design ...
So what does the research say?

The effectiveness of machine translation in foreign language education: a systematic review and meta-analysis

Sangmin-Michelle Lee

(2021)

Thirty Years of Machine Translation in Language Teaching and Learning: A Review of the Literature

JASON R. JOLLEY

Missouri State University
E-mail: jasonjolley@missouristate.edu

LUCIANE MAIMONE

Missouri State University
E-mail: lucianemaimone@missouristate.edu

(2022)
Research on MT and language learning

>100 papers since Corness (1986) cited in Jolley and Maimone (2022)

Some reviews exclude more recent sources concerned specifically with professional translator training or post-editing studies not conducted within an L2 or L3 language learning scenario.

Typical foci:
- self-reported use of MT (extent, particular uses)
- teachers’ and learners’ perception, opinions, beliefs, attitudes regarding MT
- use of MT as academic dishonesty

*empirical studies on the effects of MT use in L2 learning are still scarce*

Jolley and Maimone (2022: 27)

Much of the literature is comprised of argumentative essays and descriptions of classroom activities.
**MT in Language Teaching and Learning.** This strand comprises essays, activity reports, and empirical studies that specifically address MT use in formal learning contexts. They discuss current and potential uses of MT to develop language skills (primarily translation and L2 writing), describe MT-supported activities, and report findings of experiments. Authors of these studies recommend best practices for using MT to support language learning.

Jolley and Maimone (2022: 27)
Contemporary Relevance of “Older” Sources?

E.g. Somers (2003) and even Enkin & Mejías-Bikandi (2016) propose activities that could enhance **metalinguistic awareness:**

*The ability to focus attention on language as an object in and of itself, to reflect upon language, and to evaluate it.*  
(Thomas 1988: 531, in Thue Vold 2018: 67)

General assumption is that the MT will get the “structure of interest” wrong, and that students can comment on this.
Examples of difficult sentences from Somers (2003: 231) and Enkin and Mejías-Bikandi (2016)

- L'oiseau entrait dans la chambre.
  - The bird entered the room.
- On a dormi dans ce lit.
  - We slept in this bed.
- Nous venons de finir de lire ce livre.
  - We have just finished reading this book.
- Ich habe grossen Hunger.
  - I'm very hungry.
- Es wird getanzt und gegessen.
  - There is dancing and eating.
- Hans will, dass Kurt sein Frühstück isst.
  - Hans wants Kurt to eat his breakfast.

Queremos que nos digan la verdad.
- We want them to tell us the truth.

Es importante que nos digan la verdad.
- It is important that they tell us the truth.
Sample structures of interest (Enkin and Mejías-Bikandi 2016)

I was told there is absolutely nothing they could help me with...

GT 2014 & 2015: Me dijeron que no hay absolutamente nada que me podía ayudar con...

HT 2016: Me dijeron que no hay absolutamente nada con lo que me pudieran ayudar...

GT 2022: Me dijeron que no había absolutamente nada con lo que pudieran ayudarme...
Empirical Study of Metalinguistic Awareness in MT Use  (Thue Vold 2018)

learners of French as an L3 in Norwegian upper secondary school
read two different machine-translated versions of the same text (Google Translate & Bing Translator)
decided which MT was better and explained why

Thue Vold concludes that while the use of MT texts to develop learners’ metalinguistic awareness has “considerable potential”, “training, scaffolding techniques and guidance from the teacher are of paramount importance” (ibid.: 89) as, left to their own devices, learners may not explore fruitful avenues of analysis, and their group conversation may even reinforce misconceptions about language (ibid.).
MT in writing tasks (O’Neill 2019)

>300 American students wrote short compositions in French and Spanish using:
- Google Translate with prior training
- Google Translate without prior training
- an online dictionary with prior training
- an online dictionary without prior training
- or no technical aid at all.

O’Neill found that:
- the GT+prior training group scored better than all others on their compositions, followed by
- online dictionary+prior training group.
- In post-tests conducted one week later, and three to four weeks later, where students no longer had access to the tool in question, the Google Translate + Training group no longer performed better than the other groups.
- The advantage of using the tool was short-lived and dependent on the continued availability of the tool.
MT in writing tasks (Fredholm 2019)

Fredholm (2019) tracked lexical diversity in compositions written over the course of a full school year by 31 Swedish upper secondary school pupils of Spanish as a foreign language.

- ~50% used a printed dictionary as a translation tool
- ~50% used Google Translate

Use of MT was associated with higher lexical diversity, and hence better performance, as long as students continued to have access to the tool, but once access was removed, the effect vanished. Again, the benefit bestowed by use of MT seemed dependent on the continued availability of the tool.
MT and Syntactic Priming  (Resende & Way 2021)

Brazilian Portuguese speakers exposed to Google Translate’s correct use of particular structures reproduce those correct structures, even after a delay of 24 hours.

Evidence of implicit learning about syntax from NMT outputs in some of the participants.
Some conclusions from existing research

- Language learners use MT and rather than trying to outlaw its use, it is better to take a nuanced approach, based on an understanding of where MT can be more or less helpful, depending, perhaps, on the extent and context of use.
- Language learners make better use of MT when they have received appropriate training.
- Language learners can generally benefit more from MT if they already have reasonably good proficiency in the foreign language (O’Neill 2012, Resende & Way 2021).
Other Observations: Protocols for conducting research with FOMT?

Figure 1: input to Google Translate with sentence-initial lower case and no sentence-ending punctuation, produced 19 October 2021

Figure 2: input to Google Translate with sentence-initial upper case and sentence-ending punctuation, produced 19 October 2021

cf. Ducar and Schocket (2018: 785)
It’s great when researchers:

- make clear the extent to which, e.g., performance in writing tasks or evidence of increased metalinguistic awareness are ends in themselves, or are assumed to be indicative of language development over time
- give the technology a fair chance
- situate the use of tools like MT within principled approaches to SLA/language learning

*before talking about the use of technology in language classrooms, we must talk about how additional languages are learned*

(Egbert, Chao and Hanson-Smith 1999: 2)
Acknowledgments

Thanks to the members of the MultiTraiNMT Team.

This work is supported by the European Union’s Erasmus+ Strategic Partnership Programme, as part of the project known as “MultiTraiNMT - Machine Translation training for multilingual citizens” (project ID: 2019-1-ES01-KA203-064245).
MultitraiNMT invites colleagues from language learning and translation to become associate partners of the project in order to evaluate and share the project outcomes as well as other machine translation resources. You are welcome to join us!

dorothy.Kenny@dcu.ie
Website: multitrainmt.eu
Twitter: @multinmt

From January 2022: Try the MutNMT platform!

https://ntradumatica.uab.cat/
References


