

This work is licensed
under a Creative
Commons Attribution
4.0 International
License.

Teachers' perceptions of machine translation as a pedagogical tool

Saziye Tasmedir

s.savaskan2@newcastle.ac.uk

Elaine Lopez

elaine.lopez@newcastle.ac.uk

Müge Satar

muge.satar@newcastle.ac.uk

Nick G. Riches

nick.riches@newcastle.ac.uk

Newcastle University, UK

The role of Machine Translation (MT) within the L2 classroom has proven controversial. While many students value MT to support various learning objectives, teachers fear its potential to promote passive learning, and undermine academic integrity. Nonetheless, improved reliability and accessibility of MT, combined with student attitudes, makes it increasingly hard to “Google-proof” the classroom. Alternatively, MT might be incorporated into the classroom, while highlighting to students its limitations. A key factor in this dynamic is the attitude of classroom teachers. The current study details the development of a pilot translation app, Transpose, which then acted as discussion prompt to explore teachers' attitudes towards MT in focus groups. Participants were four UK language teachers working with various age groups and in various contexts. The Transpose app exploited possible advantages of MT by visually comparing the syntactic structure of source and target languages. A thematic analysis found that, overall, teachers viewed translation as an important pedagogical tool. Regarding MT, views were mixed. They enthusiastically suggested novel uses, such as supporting children with English as an Additional Language, yet also felt that such technology was best employed to support independent study in advanced learners. Technological and institutional factors still prevent the seamless integration of IT-related methods within classroom contexts. Overall, while teachers expressed positive attitudes towards translation, and a cautious

acceptance of the place the Transpose app could take within the classroom, there remain attitudinal barriers to widespread adoption.

Keywords: translation, machine translation, pedagogy, syntax, vocabulary

Introduction

The potential benefits and drawbacks of translation between the L1 and L2 is controversial within language teaching. While, traditionally, teachers have discouraged translation, fast, accessible, and increasingly accurate Machine Translation (MT) makes this increasingly difficult. “Google-proofing” the classroom may no longer be a realistic option. An alternative is to exploit the potential benefits of this emerging technology, while drawing attention to its pitfalls. The current paper details the creation of a prototype app which harnesses Google Translate (GT) to demonstrate cross-linguistic similarities and differences in sentence structure. The app was then used as the basis for two focus groups with classroom language teachers to explore the role of MT in the classroom, and teachers’ attitudes towards it. Given the socially constructed nature of technological innovation, the opinions of potential end-users regarding how they might exploit the app in class, and MT more generally, were explored. In the following sections we discuss the advantages and disadvantages of classroom-based translation, the role of MT and teachers’ attitudes towards this technology. We then describe the development of the translation app, and the findings of a focus group wherein teachers discuss its benefits, drawbacks, and possible classroom applications.

Background

The role of translation in L2 learning

The role of translation in L2 pedagogy, either to or from the target language, is much disputed. Negative attitudes stem from the historic shortcomings of the Grammar Translation approach, and the strong influence of the Communicative Approach which discourages translation in the classroom, or use of the L1 as a metalanguage, for example, to explain grammatical rules or provide instructions for activities (Machida, 2011). Nonetheless, despite the continued influence of the Communicative Approach, views of translation are relatively diverse. For example, some cultures and societies adopt a positive view of translation such as the *yakudoku* method in Japan (Nishino, 2008). Additionally, the decline in dominance of strictly-prescribed methods (Hinkel, 2006) means that teachers have more leeway to utilise translation. Mixed attitudes to translation are evident in a 2013 survey of 824 mainly-EU-based teachers, which involved ranking classroom activities (Pym et al., 2013). Translation was the most controversial activity leading to a wide range of both positive and negative opinions.

Empirical studies demonstrate the pitfalls of direct translation. Lexical

errors may reflect false cognates (false friends). For example, an L1 Spanish learner of English may describe someone as “educated” meaning “polite” because the Spanish word for “polite” is “*educada/o*.” Grammatical errors may also reflect direct translation. For example, an L1 Spanish learner of English may transfer reflexive particles from their L1, e.g. “He weared **him** a dress” (Alonso & Rosa, 1997). The reflexive particle of *llevarse* [wear] is erroneously transferred. Direct translation errors may also result in stylistically inappropriate language. For example, an English learner of Spanish might overuse the periphrastic passive, which is less frequent in Spanish than English. There is also evidence that direct translation impedes accuracy and fluency. Research in “thinking for speaking” (Slobin, 1996) has found that speakers of different languages must plan utterances differently. While an English speaker may combine manner and path on the verb, for example, “the ball **rolled down** the hill,” Spanish allows manner information to be included only via a sentence-final present participle, or extra-linguistic means (e.g., gesture). Dependence on direct translation may hinder the development of language-specific sentence planning mechanisms.

However, despite the pitfalls of direct translation, early-stage learners have a tendency to mentally translate. According to Cohen et al. (2000) 80% of students performing an L2 writing task reported mental translation from the L1. Furthermore, experimental studies suggest that, even in fluent bilinguals, the non-target language is highly activated. When performing a naming task in one language, performance is significantly faster when there is a cognate in the non-target language (Costa et al., 1999). Consequently, cross-linguistic interaction and competition is a natural outcome of having knowledge of multiple languages. Therefore, an enforced separation between the L1 and the target language may be misguided. According to Widdowson, “while in the classroom the teachers try to keep the two languages separate, the learners in their own minds keep the two in contact” (Widdowson, 2003, p. 150, as cited by Cook, 2003).

One argument for translation in the early stages is that, given learners’ instinctive dependence on translation, a key task is learning when *not* to directly translate. A means of addressing this explicitly is to directly compare translation-equivalent sentences across languages. Indeed, the teaching of false cognates to avoid lexical translation errors is well established, for example dictionaries of so-called “false friends” (Hill, 1982). By contrast, techniques which focus on direct translation errors in syntax are less common. However, such a process may be beneficial to overcome translation errors. Moreover, to develop appropriate “thinking-for-speaking,” an important first stage is observing when L1 “thinking-for-speaking” strategies cannot be transferred to the target language.

In drawing attention to syntactic divergence between the L1 and the target language, translation improves our structural understanding of the latter, and is therefore consistent with recent pedagogical trends involving a Focus on Form (Machida, 2011). This argument is also made by Cook (2007) in an article which argues that the learning potential of translation has been overlooked,

and that it may play an important role in the L2 classroom. Moreover, Machida (2011) argues that translation fosters understanding of how culture and context may impact on language use and meaning. For example, the translatability of idioms across languages can promote discussions related to cultural similarities and differences. Moreover, words which initially appear directly translatable may have different connotations in each language. For example, German *Kamel* often denotes stupidity / stubbornness, like English *donkey*, whereas Spanish *camello* can refer to drug pushers (Chamizo Domínguez & Nerlich, 2002).

The impact of machine translation

The emergence of MT has inevitably led language teachers to reassess the role of translation in the classroom. Recently, GT translations have greatly improved in terms of fidelity to the source material, grammaticality, coherence, and resemblance to native speaker output. For example, a study employing the GT 2016 version (Tsai, 2019) found that Taiwanese university students of English produced better essays when translating automatically from the L1 versus writing directly in the L2, with quality calculated using automated measures of error frequency and lexical density. Likewise a study of Hong Kong primary children (Stapleton & Kin, 2019) found that their English teachers could not distinguish between directly-written and Google-translated essays, regarding the latter as containing marginally-better syntax and vocabulary. However, despite this progress, GT and other MT technologies, still faced challenges, including use of register and grammatical aspect, as recently as 2018 (Ducar & Schocket, 2018).

Improvements in MT, combined with improved accessibility via handheld devices, have important ramifications for pedagogical practice. For students, MT provides an accessible tool for purposes ranging from the benign, like comprehending text, to the dubious, such as generating assessed essays. Students in higher education settings view MT positively, for example as a method to expand vocabulary (Clifford et al., 2013; Niño, 2009). In a study investigating Korean university students' use of MT to check their English essays (Lee, 2020) students felt that it facilitated the use of context-appropriate words. By contrast, teachers are concerned with threats to academic integrity, such as generating L2 essays from the L1, its impact on student motivation, and also its capacity to displace traditional classroom approaches (Clifford et al., 2013; Crossley, 2018; Ducar & Schocket, 2018; Lee, 2020). Stapleton and Kin (2019) assessed the accuracy of GT on the writings of Hong Kong primary school pupils, and all twelve primary teachers who participated in the study disagreed with the use of MT to translate above the word level when interviewed. The disconnect between students' and teachers' views is striking (Clifford et al., 2013).

Some teachers have prohibited MT in the classroom, while others have adopted a flexible approach, acknowledging the potential benefits and allowing limited use whilst simultaneously highlighting pitfalls (Clifford et al., 2013). Ducar and Schocket (2018), argue that "google-proofing" a classroom

is a fruitless endeavour, and teachers should therefore make classrooms “google-irrelevant,” for example, by creating activities where students have little motivation to exploit MT. They argue that teachers must engage with MT, discussing its pros and cons, and under what circumstances it can undermine academic honesty. Finally, teachers might encourage students to consult other electronic resources such as dictionaries, corpora, or search engines to reduce over-dependence on MT.

The focus of the current study

Despite the controversy surrounding MT, it nonetheless has pedagogical potential for learning L2 lexis and sentence structure. Moreover, because it works automatically on any text, it enables translation activities at scale with minimal teacher preparation. Students also regard it favourably. However, teachers are more circumspect, expressing concerns over academic integrity and student motivation, acquiring errors from faulty translations, and threats to the traditional role of the classroom teacher.

The current study explores the potential use of a GT-based web app as a means of learning the target grammar. The newly developed Transpose app highlights syntactic differences between the L1 and the target language. The app was used as a prompt in a focus group with classroom teachers, which led to an evaluation of the app and facilitated the exploration of attitudes to MT more generally. Technology use is shaped by key social groups (e.g. teachers and students), and the meanings that they invest in a particular tool (Jones & Bissell, 2011; Kline & Pinch, 1996). An example of this process is the occasional re-engineering of technologies for purposes beyond those originally intended (Jones & Bissell, 2011). Within the MT literature there is also evidence for such re-engineering, such as the use of MT for electronic “peer-editing” (Lee, 2020). Whereas in standard peer-editing, feedback is obtained from classroom peers, in an MT context, the “peer” is the translation platform itself. By presenting the app within focus groups we investigated how potential end users might utilise it. This approach is consistent with Edwards & Holland’s (2013, p. 53) view that “researchers can utilize a range of other textual, visual and creative tools to engage interviewees and stimulate discussion as part of qualitative interviews, and thereby reveal aspects ... interviews using words alone cannot reach.”

Methods

The current research involved the development of an app, Transpose, which was presented to users via two consecutive qualitative focus group interviews comprising of four language teachers. The aim was to appraise Transpose as a potential learning aid and further explore teachers’ perceptions of MT as a pedagogical tool. We used an exploratory qualitative methodology (Silverman, 2013) to obtain rich descriptions of the participants’ reflections.

Development of the app

This used GT as the translation engine, and visually manipulated the output to highlight the syntactic structure of the sentences. GT was embedded within a web app created using the Shiny Apps framework (Chang et al., 2020) a library written for the open-source language R (*R: A Language Environment for Statistical Computing*, 2020).

The app asks the user to enter textual data and attempts to determine the source language. This may be overridden when language detection fails. Users then click a tab displaying the source and target texts. The structure of the text is highlighted in numerous ways. Firstly, source and target languages are presented vertically, encouraging the user to engage in direct translation, such as mapping a preposition in the source language to its translated equivalent in the target language. Secondly, words are highlighted according to their syntactic classes. In the default colour scheme, verbs and verb-related items (e.g., adverbs, auxiliary verbs) are highlighted in shades of pink, nouns and noun-related items (e.g., determiners, adjectives) are highlighted in shades of blue, prepositions are yellow, pronouns are green, conjunctions are orange, and “social” words, such as greetings and exclamations, are brown. This colour scheme supports syntactic analysis firstly by highlighting words within grammatical phrases the same colour; shades of pink for the verb complex (verb and associated words), and shades of blue for the noun phrase, and secondly, by using brighter colours for words playing important grammatical roles. Verbs, which determine the number of arguments, and also clauses (via subordination,) are coloured deep pink, while subordinating conjunctions and prepositions are coloured orange and yellow respectively. As GT itself does not identify the syntactic classes, the package *udpipe* (Wijffels, 2020) was used for this purpose. This exploits parsing models developed by the Universal Dependencies international project (Nivre et al., 2019). This aims to develop “cross-linguistically consistent treebank annotation for many languages, with the goal of facilitating multilingual parser development, cross-lingual learning, and parsing research from a language typology perspective” (<https://universaldependencies.org/introduction.html>). Finally, the user could identify word classes by hovering the mouse over the text. The app was programmed by the last author, and the technique of colouring words initially exploited in an app created for Speech and Language Therapy students to analyse the sentence structure of child language (<https://github.com/nickriches/transpose>).

Unfortunately, parsing (tagging of word classes) is only around 95% accurate. Moreover, the languages covered by GT and Universal Dependencies do not perfectly overlap. A total of 57 out of the 104 languages covered by GT (55%) have no Universal Dependency model. For these languages, the translation will work, but word classes cannot be coloured. Languages with no Universal Dependency model tend to be non-European with comparatively few speakers.

Screenshots of the app are provided in Figure 1. The app may be downloaded from (<https://github.com/nickriches/transpose>). A working copy is available for evaluation purposes, and instructions are provided on how to install

it on your own server. Some technical knowledge is required for this, and knowledge of the R computing framework is advantageous.

Entering text

Translation App Instructions Let's get started! ▾ Let's ex

How do you wish to enter your data?

☐ Upload file (.doc, .docx, or .txt)

☒ Enter text in textbox

Enter text here...

Tardigrades, known colloquially as water bears or moss piglets, are a phylum of eight-legged segmented micro-animals. They were first described by the German zoologist Johann August Ephraim Goeze in 1773, who called them little water bears. In 1777, the Italian biologist Lazzaro Spallanzani named them Tardigrada, which means "slow steppers".

Choosing the target language

Translation App Instructions Let's get started! ▾ Let's explore! Colours Punctuation

Text

(up to 1000th character)

Tardigrades, known colloquially as water bears or moss piglets, are a phylum of eight-legged segmented micro-animals. They were first described by the German zoologist Johann August Ephraim Goeze in 1773, who called them little water bears. In 1777, the Italian biologist Lazzaro Spallanzani named them Tardigrada, which means "slow steppers".

from... to...

English Spanish; Castilian |

Afrikaans
Albanian
Amharic
Arabic
Armenian
Azerbaijani
Basque
Belarusian

The translated output

(NB VERB label is shown via mouse hovering)

Translation App Instructions Let's get started! ▾ Let's explore! Colours Punctuation

Table will take a few seconds to appear/refresh...

Search:

sentence_id	sentence_coloured
All	All
	VERB
1	Tardigrades , known colloquially as water bears or moss piglets , are a phylum of eight - legged segmented micro-animals .
1	Los tardigrados , conocidos coloquialmente como osos de agua o lechones musgosos , son un filo de microanimales segmentados de ocho patas .
2	They were first described by the German zoologist Johann August Ephraim Goeze in 1773 , who called them little water bears .
2	Fueron descritos por primera vez por el zoólogo alemán Johann August Ephraim Goeze en 1773 , quien los llamó pequeños osos de agua .
3	In 1777 , the Italian biologist Lazzaro Spallanzani named them Tardigrada , which means " slow steppers " .
3	En 1777 , el biólogo italiano Lazzaro Spallanzani los llamó Tardigrada , que significa " steppers lentos " .

Showing 1 to 6 of 6 entries

Figure 1. Screenshots of the “transpose” translation app

Participants for the two focus group interviews were recruited through regional networks, including direct emails to schools in the local area, a local language learning association newsletter, and local events. Participants were offered the opportunity to learn about a new MT web app and were also offered refreshments as incentives for participation.

Focus group participants were key informants representing their own and their colleagues’ attitudes towards MT within their specific contexts. These contexts were a state middle school (Spanish and French), a state secondary school (French and German), and a community school (Greek) in the Northeast of the UK. Information about each participant and their respective teaching context is presented in Table 1.

Table 1. Participant background information

Pseudonym	School	Languages/courses taught	Age group taught
Sophie	Secondary school	GSCE French GSCE German	11–16
Natalie	Community school	Greek	Adults
Emma	Middle school	French Spanish	9–13
Barbara	Middle school	French Spanish	9–13

Learners at the middle school were aged 9–13, largely focusing on listening and speaking skills with few written materials. The teachers in this school do not use textbooks and display lesson content on interactive whiteboards with many games. They heavily rely on technology for whole class work, yet individual access to technology is limited as learners cannot use their own devices (e.g., mobile phones) and institutional devices require substantial time investment due to institutional booking systems and technical issues in class, such as password problems. Teaching provision is 1 hour per week, and learners do not need to prepare for an exam. Thus, the focus is on making language-learning enjoyable. The school offers French to years 5 and 6 (age 9–11). The learners then take Spanish as the second foreign language in years 7 and 8 (age 11–13).

Learners at the secondary school were aged 11–16 and were studying towards the national UK General Certificate of Secondary Education (GCSE) qualification in French and German, taken at 16. Teaching was exam-oriented, and teachers felt that supporting learners to pass the GCSE was a large part of their role. The GCSE tests all four language skills (reading, writing, listening, speaking) via open-ended questions. The writing component includes a section on translating sentences from English to the target language, and vice versa. The school does not have tablets but has bookable computer suites. The students largely come from deprived areas and do not have access to technologies or sufficient internet data packages for online activities outside of class. The school has a large proportion of multilingual students, especially from

migrant families, who tend to make comparisons across their languages. While the school has a no-mobile-phone policy, the use of learners' personal mobile phones is partially tolerated for teaching and learning.

Learners at the community school were predominantly adults learning Greek at evening classes. It does not have a specific building, but instead rents space from mainstream schools. Consequently, there is no readily available classroom technology or any form of IT support, meaning the teachers must provide the technology themselves. Despite the limited resources, teachers in this school have a positive approach towards incorporating technology into their classes. As the learners are generally adults, they tend to bring their own devices, such as mobile phones and tablets, to class and are encouraged to use them during class time to support learning and teaching.

Procedures, and methods of data collection and analysis

Data collection was conducted twice. The first focus group (Sophie and Natalie) took place in November 2019. Initial data analysis and app revision was conducted before the second focus group (Emma and Barbara) in February 2020. Both focus groups generated 137 minutes of audio-recorded data.

Data collection tools included (1) a brief background questionnaire, (2) a semi-structured interview guide, and (3) a demonstration and discussion of the new MT web app. The background questionnaire allowed researchers to collect demographics about the teachers, such as the languages they teach, teaching experience, gender, and qualifications. The semi-structured questions elicited teachers' thoughts and attitudes in three areas: use of technology in language learning and teaching; use of L1 and translation in language learning and teaching; and the role of teaching sentence structure and word class in their teaching practice. The MT web app Transpose (<https://github.com/nickriches/transpose>) was demonstrated as an interactive workshop. The researchers then elicited an appraisal of the app regarding potential teaching applications and areas for improvement. This facilitated app development while acting as an interviewing tool to trigger further reflection based on a concrete artifact (Edwards & Holland, 2013).

The focus group data were transcribed and imported into NVivo 11. Data were analysed by two researchers using Braun & Clarke's (2006) six-phase inductive thematic analysis procedure. This involved familiarising oneself with the data through transcription and re-reading, initial descriptive coding, collating codes into related thematic categories, formulating and refining the themes to eliminate contradictions and overlaps, generating clear definitions of the final themes and sub-themes, and finally writing a report with relevant data extracts. Data extracts presented below were chosen based on their representativeness, succinctness, and effectiveness in illustrating the theme. During this process, individual perspectives were discussed, and final themes were determined unanimously. Within this process, constant comparison (Richards, 2003) of new data with coded data enabled generation of themes progressively and ensured trustworthiness. The following major themes were identified: barriers

to MT adoption in teaching, motivations towards MT adoption and feasibility of adopting the web app, Transpose, in teaching.

Findings

Here we outline major themes regarding teachers’ comments, emphasising relevant relationships between the themes or factors specific to a given context. The first main theme which emerged was perceptions of barriers to MT adoption in language teaching (Figure 2).

Barriers to MT adoption in language teaching

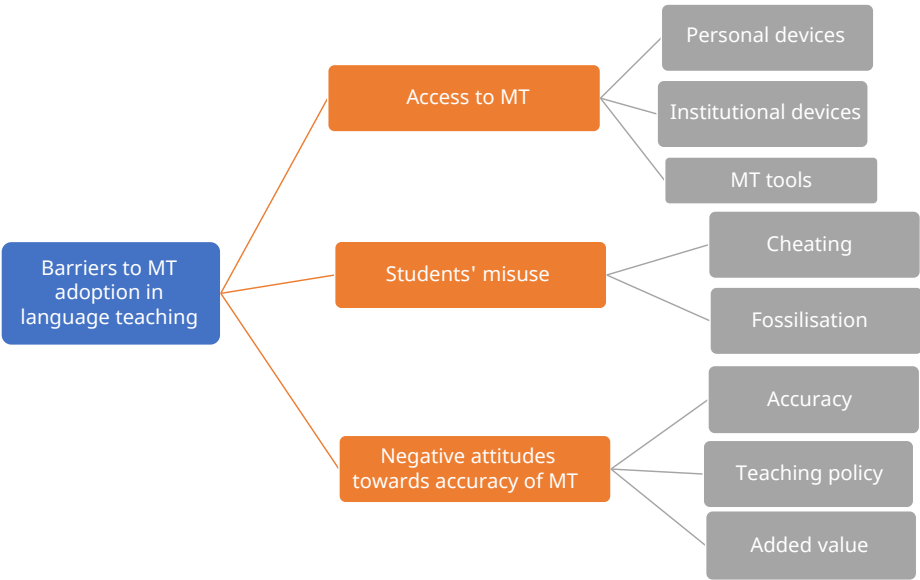


Figure 2. Theme 1: Barriers to MT adoption in language teaching

Barriers were grouped under three sub-themes. First, access to technology was a key issue in the state middle and secondary schools. The middle school had institutional mobile devices that could be borrowed, but teachers could encounter several time-consuming challenges (e.g. passwords not working), which made the use of such devices worthwhile only if the whole lesson was structured around technology use. According to Barbara, if there were fewer challenges, “It would be great just to do it, do a little bit of IT.” In the secondary school, there were no mobile devices, but only computer-clusters, which created similar challenges, whilst the community school had no institutional technology support. Thus, in the community school, any technology-related activity depended on learners’ personal devices. Natalie welcomed their use with her adult students.

However, in the middle and secondary school contexts, there were also institutional policies against use of personal devices in class with a “no-mobile-phone-policy” (Sophie). This was motivated largely by concerns over access to inappropriate content on mobile phones and vulnerability to viruses. Though institutional devices, such as school tablets, tracked a history of access, teachers

felt overwhelmed by the responsibility of policing the students' internet use. Emma described this situation as follows:

And we still have problems with the tablets like I know like some kids have had to like be spoken to because in order to even like in the classroom, how they've taken inappropriate pictures with them or they've been on sites that weren't blocked and they shouldn't have been on. Obviously we've got the history on them. But even Yeah, even the teachers who do use it regularly still come across problems. So it's stuff like that that puts you off using them.

In the secondary school, there was also a no-mobile-phone policy, but it was not as strict. There were no institutional mobile devices, and provided that mobile phones were employed for learning and teaching activities that the teacher can justify, their use could be sanctioned. As Sophie put it: "policy is just that students, unless the teacher has a reason for it and justifies it, are not allowed to get their phones out in class ... it's a bit of a grey area at the minute".

A final issue in relation to access to MT, specifically, was institutional policies against use of MT. Although all the teachers in our focus groups found translation to be a beneficial activity in language learning, none used MT, and in the middle school GT was banned because "Like if they'd have to do homework and things like that, and you ask them to do like a little bit of writing, you can immediately tell who's put it into a translator" (Emma).

The second barrier towards adoption of MT was concerns around students' misuse of the technology and becoming "reliant on it" (Sophie). At both middle and secondary school levels, students' cheating on their homework was a concern since it was believed that students were "not learning anything really" (Emma) by copying and pasting translated structures. This was such a widespread practice that it caused some teachers to avoid writing homework. The teachers could tell work was translated either because it involved untaught structures (Sophie) or it did not make much sense (Barbara). Teachers were also concerned that inaccurate, poor translations could lead to fossilisation of the L2, especially when it becomes habitual and goes undetected (Natalie).

Finally, teachers' negative attitudes towards MT were a barrier towards its adoption. Emma from the middle school was perhaps the most sceptical participant, commenting on the unreliability of MT, and the need to check MT output "on three different websites". She also criticised MT for lack of register information or regional use (e.g., American Spanish), which was available on other online reference websites. Given these concerns, coupled with challenges surrounding use of technology in the class, she questioned the added value of MT: "Is it, is it worth the effort?"

Motivations towards MT adoption in language teaching

This main theme involved three sub-themes: positive attitudes towards the role of technology, the role of translation and affordances of MT in the language classroom (Figure 3).

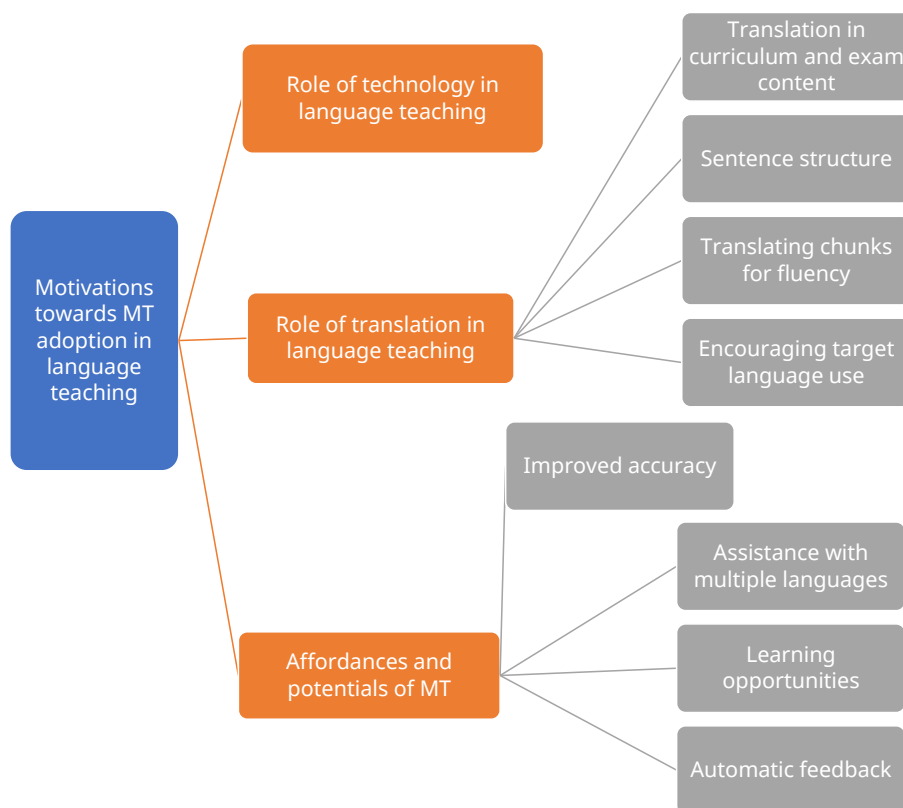


Figure 3. Theme 2: Motivations towards MT adoption in language teaching

Overall, participants had extremely positive attitudes towards technology in the language classroom. Emma and Barbara described interactive whiteboards to be “indispensable” with all teaching content being available and presented electronically, which is the norm for the students. As Barbara commented: “I don’t think I’ve used a textbook for years ... I don’t think they [students] expect it.” This comment related to whole group teaching with available technology in the classroom. When integration of technology was less seamless, for example teachers having to borrow institutional mobile devices for individual access, technical challenges caused “disruption” especially given limited classroom time for language teaching (1 hour per week). They valued teacher-student and student-student interaction above student-computer interaction. In the secondary and community schools, no negative aspects were mentioned, and the motivational aspects of technology were emphasised, such as providing immediacy and “instant gratification ... anything that sort of hooks them in” (Sophie).

In all three contexts, participants also expressed positive views on translation in language learning and teaching. This was influenced by the existence of translation activities within the curriculum and exam content, and teachers’ perceptions towards translation as a useful method to teach sentence structure, to improve fluency through translating chunks, and to encourage further target language use in the classroom.

Especially in the state middle and secondary schools, translation was

perceived as an important element of teaching since it is tested at GCSE level and the curriculum includes translation activities. In the middle school, translation activities predominantly targeted the word and phrase level by eliciting the L1 and L2 equivalent of words and phrases. The secondary school teacher noted that the translation component of the GCSE was “quite a big part of the exam even at the foundation stage. So we have to teach translation skills” (Sophie). Although at middle school, the teachers did not need to engage with exam preparation per se, they still felt translation activities were important since “it’s the skills that they need when they get up to a higher level.” (Barbara) and “you kind of like prep them in different lessons” (Emma).

All teachers agreed that students struggle with metalanguage for grammar because English grammar is not taught explicitly. They believed language comparison between L1 and L2, but also L2 and L3 (in the middle school where both French and Spanish were taught) was a useful method for understanding structure and reinforcing recall. As Barbara said: “it’s quite nice that we can in a French lesson just say well in Spanish you do it like this, can you remember. But, and then it might stick in their minds a bit more.” The teachers have also developed certain translation-based activities. For instance, Sophie independently developed a technique which she called “dodgy English.” Students divide a sentence into chunks, first with a direct translation into a grammatically incorrect form (e.g. *J’ai les cheveux bruns* → *I have the hairs brown*) and then changing into correct English (*I have brown hair*). In the middle school, the teachers also used L2 (French) and L3 (Spanish) translations to help learners understand similar grammatical phenomena such as “the position of adjectives” (Emma). At middle school level, since the children were younger, most translation activities involved games. One such game was a dice game, which Emma explained as follows:

they might have like, six, either six sentences in English or six pictures, and they roll the dice and whatever it lands on, they have to turn that into the target language. If their partner is happy with what they’ve said they tick it off. If their partner spots a mistake they’re not allowed to tick it off.

The teachers also perceived translation to be part of second language fluency development and a cognitive process employed especially at early developmental stages of speaking and particularly at phrase level. Translation of chunks and phrases was perceived as a way to think in the target language and become fluent. The secondary school teacher, Sophie, observed that the introduction of translation into GCSEs, and the consequent focus on translation in the classroom, has coincided with better listening skills and fluency. In the community school where learners were largely beginner level students, Natalie allowed her students to speak in English, but also encouraged them to say small chunks in Greek. This was echoed by teachers at the other two schools where translation was used to encourage further target language use in the classroom.

The third sub-theme within motivations for MT adoption was teachers’ positive perceptions towards the affordances and potentials of MT technologies, which they largely expressed following appraisal of the web app: Transpose.

First, participants were largely impressed with recent developments in the accuracy of online translation based on their trial of the app for French, German and Greek, finding the app outputs to be “really good!” (Natalie). Natalie perceived MT to be especially beneficial for bilingual and multilingual students since the community school students had various L1 backgrounds and teachers struggled to assist students from diverse L1 backgrounds with comparative analysis. As Natalie said: “I have [students] that know other languages like German ... so they’re trying to do the comparison themselves ... but I just know English and Greek and some French so I’m not very good at it. But if an app could do that hehe like for me.” MT in general and Transpose specifically were also perceived as valuable for immigrant children in middle and secondary schools, both to enable L1 comparisons, but also to support them widely in the classroom, for example when following classroom instructions (Sophie). Sophie also thought Transpose could help her understand the learners’ L1 structure, thereby enabling her to offer corrective feedback and better understand mistakes due to L1 interference.

When combined with part-of-speech tagging, as the Transpose app does, MT was thought to create several learning opportunities. Sophie was particularly keen and suggested that the app can make the students’ “written and translation work much more accurate.” She also thought that parsing allows users to harness MT affordances for actual learning. Instead of just copying and pasting, students can play with changing parts of the sentence and observe how this affects the translation. According to Sophie, Transpose highlights language learning affordances through comparison of sentence structure and provides “a way of harnessing Google to actually teach them how the languages work.” For Sophie, this opportunity could address “a big controversy” within the “languages teaching community” whereby teachers try “to get away from kids using Google translate”, and thus constitutes a means of shifting negative attitudes among teachers regarding misuse of MT.

In contrast, the middle school teachers were more cautious towards MT in the classroom, suggesting supervised use with “prior checked” language content, rather than more flexible autonomous use of MT:

I would probably tend to use it like we were saying today so the kids doing like the translation task or whatever. So rather than go through it all as a class if they did have the tablets or something in pairs to translate on the whiteboards and then check the right answer. Obviously, we would have to look at it ourselves to make sure it’s given the right option. (Emma)

Feasibility of adopting the MT web app, Transpose, in teaching

The final theme which emerged from the focus groups relates to the Transpose app itself, and the subthemes are shown in Figure 4.

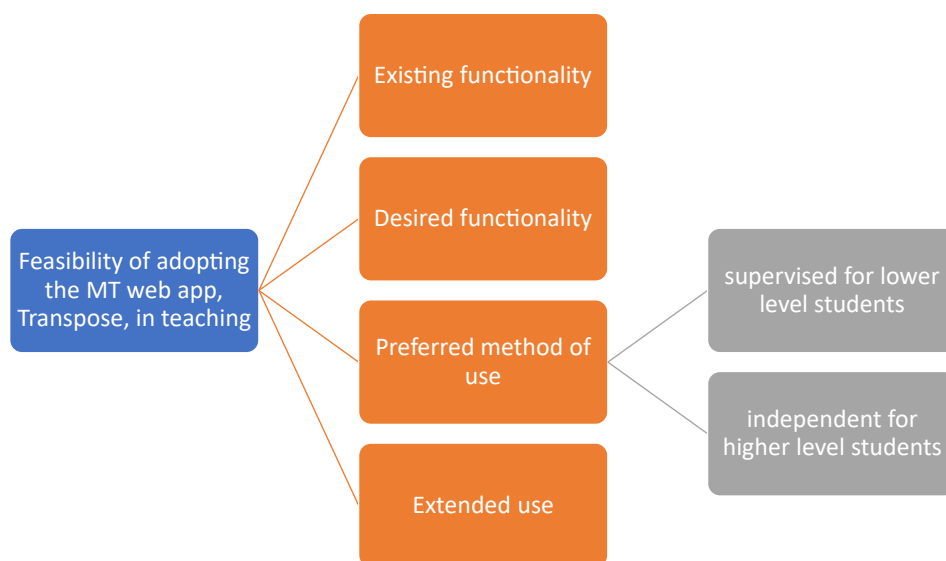


Figure 4. Theme 3: Feasibility of adopting the MT web app, Transpose, in teaching

As mentioned above, there was general enthusiasm for the use of MT and the Transpose app to demonstrate language structures and support multilingual students especially at the secondary and community schools. The teachers were keen on the ability to analyse sentence structure using colours and compare languages by visualising differences. One participant felt that this would be particularly useful for students learning English as an Additional Language (EAL) if translation were possible from their first language (Natalie). One teacher (Sophie) had already employed her own method for comparing language structure, as explained in the previous section. This is analogous to the way the app works.

In terms of existing functionality, the web-based nature of the app was appreciated as it can be used on mobile phones and tablets, unlike locally installed apps (as mentioned in Theme 1). Another feature that was mentioned as a positive was line numbering. The teachers also made some specific suggestions for improvements of Transpose (please see Figure 1 for a reminder of the app interface). In response, the app was developed to allow translation between more than two languages, to allow users to search for word classes, and to create bespoke sets of colours. These features are present in the latest version. The suggestions which led to some of these changes are as follows:

1. ability to correct wrong or inappropriate translations (i.e., machine learning) (Natalie),
2. ability to search for a single word class since multiple colours can appear confusing (Sophie),
3. ability to assign a specific colour for a specific word class via a dropdown menu to increase usability (e.g., “adjectives blue”),
4. the ability to save bespoke colour schemes, designed by the teacher, across sessions (Emma: “Because I suppose if you’ve got your own set of

- colours and things then if they were consistent that each time you use them the students remembered that, yeah, nouns were in one colour”)
5. comparison of more than two languages particularly for multilingual students (Barbara)
 6. improving the current accessibility of the app for schools because hosting might be “a stumbling block” as it involves technical knowledge and financial investment. There could be wider uptake if access was as easy as “just to click on this link and pay ten pounds a month or whatever” (Sophie)

Regarding preferred methods or contexts of use, there was a general consensus that the app would be more suitable for students with comparatively higher-levels of language since it might appear too complicated to lower-level students. This could be either due to learners’ insufficient knowledge of grammatical metalanguage (Natalie), experience of cognitive overload since all word classes were displayed (Emma, Barbara) or translation from L2 to L3 (Emma). Another concern for this group was the danger for fossilisation especially when used independently by learners with limited proficiency:

but if I wanted to use er like for me from a different language if I want to learn a different language, I would know because I would look it up and stuff, so I wouldn’t recommend it for the little ones ... for them to use it alone I mean because they might mix them up that’s the only reason (Natalie).

For lower-proficiency levels, translation of “short sentences” (Emma) and more teacher controlled input was suggested: “if I have already prepare what I’m gonna show so there will not be any mistakes” (Natalie). Relatedly, Transpose was found more amenable for independent, unsupervised use by advanced students (such as “A level and higher GCSE”) who could double-check the information using various sources (Sophie).

In the focus group with Sophie and Natalie, there was a discussion of using Transpose beyond the state and community school systems. The teachers suggested that Transpose would be welcomed by degree level (higher education) students, especially those studying modern languages; international students in the UK to improve their English; and relatedly EAL or English for Speakers of Other Languages (ESOL) students, that is, immigrants learning English to increase employment opportunities. Finally, Transpose was suggested as a useful tool for teachers either to learn a new language or to support their multilingual students (as in Theme 2).

Overall, teachers thought that Transpose could be flexibly integrated into whole class activities with the teacher using the app on the interactive whiteboard, or used for individual self-paced activities on personal or institutional devices with or without teacher guidance. The app was also perceived useful for autonomous, critical sentence structure and word class analysis, particularly for advanced level learners. As word-class tagging and MT technologies continue to improve in accuracy, Transpose will produce more accurate results, and therefore its potential as a self-study tool will be enhanced.

Discussion and conclusion

Though the participating teachers echoed concerns raised by teachers in previous studies, they were, overall, better-disposed to its pedagogical use. Negative views revolved around loss of learner autonomy, encouraging ineffective learning habits, and threats to academic integrity. For example, the participating teachers were concerned about students using MT for writing assignments in the target language. If students do not analyse the translated text or compare it to the source text, such a process does not promote learning. Moreover, it threatens academic integrity if students fail to acknowledge their use of MT. These themes of ineffective learning habits, and threats to integrity echo pre-existing studies (Clifford et al., 2013; Crossley, 2018; Ducar & Schocket, 2018; Lee, 2020). Other criticisms relate to lack of information regarding regional varieties, and sensitivity to register. These, again, echo concerns in the research literature (Ducar & Schocket, 2018). Nonetheless, participants expressed positive views of the app, particularly the ability to visualise structure using colours, and one teacher commented on its potential use with EAL children, presumably to facilitate comparison between English and the home language. Their engagement with the technology was reflected in their multiple suggestions for improvements, for example translation between multiple languages, and suggestions for pedagogical applications, such as manipulating the source sentence to see how it affects the target. According to one participant, the app provides “a way of harnessing Google to actually teach them how the languages work”.

The disconnect between the participants’ initial negative views of MT, and their positive views of the app is striking. Their attitude towards the app may reflect positive feelings towards translation and the role of technology in general. The former is demonstrated by their active development of translation-based techniques within the classroom, such as comparing sentence structure across languages, the so-called “dodgy English” method outlined above, the translation-based dice game and the technique of translating chunks within sentences, a practice similar to code-switching. They felt that translation may be a critical skill in the early stages of language learning, and the recent incorporation of translation into language GCSEs had improved learning. Such positive views may, partly, reflect the adoption of translation within the UK national curriculum and school examinations. However, the participants’ development of novel translation-based activities within their teaching practice indicates personal engagement with the pedagogical potential of translation independent of cultural and institutional factors. It is also possible that the app stimulated participants to think about the *potential* of MT in the classroom rather than its current usage. This is consistent with Clifford et al. (2013), who found that while only 27% of university-level language teachers viewed MT as “useful” or “somewhat useful,” a much larger percentage (54%) were positive about its future potential.

Regarding language technology, the teachers valued classroom technology, particularly interactive whiteboards, and suggested that students had come to expect this, rather than pen and paper. This positive view existed despite

technology-related issues including access to passwords, and difficulties preventing access to inappropriate materials. The combination of positive views on translation, combined with positive views on information technology in the classroom may have predisposed the teachers to view the app favourably, despite specific institutional policies prohibiting MT both at school and during homework, and also their pre-existing attitudes towards MT. Their impression that the reliability and accuracy of MT has improved may also have fostered a favourable attitude. Naturally, because the app was designed and demonstrated by the researchers themselves, this may have biased responses. Nonetheless, there was explicit engagement with the app and an appreciation of its potential.

In accordance with the social construction of technology (Jones & Bissell, 2011), teachers identified applications and drawbacks not considered during the initial development of the app. Regarding applications, they proposed using MT with EAL children. Regarding drawbacks, they identified an inability to correct erroneous output, and lack of awareness of register and variety. While Transpose was designed to highlight sentence structure with early-stage learners, the participants thought the app might most benefit autonomous study by advanced-level learners. Consequently, technologies designed with a particular goal frequently assume alternative purposes when they are applied in a real-world setting. The development of new technologies is dependent on social factors, and roll-out should be shaped by co-construction with end-users.

It should be acknowledged that there are a number of factors affecting the validity and generalisability of these findings. Firstly, regarding validity, as mentioned above, there may have been a “halo effect” whereby participants were favourably disposed towards the app, and MT more generally, because it was presented to them, in-person by the researchers. Moreover, the design of the app and its capabilities may have primed participants to consider more positive aspects of MT and its applicability in classroom settings. It is possible that more objective opinions may have been elicited without using the app as a prompt. Nonetheless, the use of graphical prompts to stimulate interviewees is relatively widespread (Edwards & Holland, 2013), and the current study extends this technique by introducing an interactive tool. It is plausible that without this specific prompt we may not have obtained such detailed observations from the participants.

Regarding generalisability, it should be acknowledged that the participants worked in very specific settings, with a very specific set of students and learning goals, and that results may not generalise to other contexts which are more frequently studied, e.g. university language classes (e.g. Clifford et al., 2013; Niño, 2009). Nonetheless, many themes are echoed across contexts. For example, the participants in the current study experienced similar concerns to teachers in higher education contexts regarding issues such as academic integrity and passive learning (Clifford et al., 2013; Crossley, 2018; Ducar & Schocket, 2018; Lee, 2020), and this may reflect commonalities across goals, for example, the need to assess a student’s fundamental abilities, and the promotion of active learning.

Despite the positive views of translation, and positive comments regarding Transpose, the role which translation, and MT in particular, plays in language

classrooms is nonetheless problematic and contested. The negative impact of MT, in particular the undermining of learner autonomy, and its threat to academic integrity, must be carefully managed. We must also bear in mind the considerable technological factors which impact on seamless integration of technology within the classroom. Nonetheless, findings indicate that classroom teachers may be receptive to MT in pedagogical contexts. This is consistent with existing literature which suggests that a minority of classroom teachers have positive view of MT (Clifford et. al, 2013). Given the accessibility and increasing reliability of this technology, pedagogical applications should be given due consideration, bearing in mind the needs and opinions of potential end-users.

Acknowledgements

We would like to extend our thanks to the participating teachers, and a grant from the Creative Multilingualism (grant number AHRC / AH/N004701/1) which made this research possible. Creative Multilingualism is supported by the Arts and Humanities Research Council, now part of the United Kingdom Research Council.

References

- Alonso, A., & Rosa, M. (1997). Language transfer: Interlingual errors in Spanish students of English as a foreign language. *Revista Alicantina de Estudios Ingleses*, 10, 7–14.
- Chamizo Domínguez, P. J., & Nerlich, B. (2002). False friends: Their origin and semantics in some selected languages. *Journal of Pragmatics*, 34(12), 1833–1849. [https://doi.org/10.1016/S0378-2166\(02\)00024-3](https://doi.org/10.1016/S0378-2166(02)00024-3)
- Chang, W., Cheng, J., Allaire, J., Xie, Y., & McPherson, J. (2020). shiny: Web Application Framework for R. <https://CRAN.R-project.org/package=shiny>
- Clifford, J., Merschel, L., & Munné, J. (2013). Surveying the Landscape: What is the Role of Machine Translation in Language Learning? @tic. *Revista d'innovació Educativa*, 0(10), 108–121. <https://doi.org/10.7203/attic.10.2228>
- Cohen, A. D., Brooks-Carson, A. W., & Jacobs-Cassuto, M. (2000). Direct vs. translated writing: What students do and the strategies they use. University of Minnesota. Center for Interdisciplinary Studies of Writing.
- Cook, G. (2007). A thing of the future: Translation in language learning. *International Journal of Applied Linguistics*, 17(3), 396–401. <https://doi.org/10.1111/j.1473-4192.2007.00160.x>
- Costa, A., Miozzo, M., & Caramazza, A. (1999). Lexical selection in bilinguals: Do words in the bilingual's two lexicons compete for selection? *Journal of Memory and Language*, 41, 365–397.
- Crossley, S. A. (2018). Technological disruption in foreign language teaching: The rise of simultaneous machine translation. *Language Teaching*, 51(4), 541–552.

- Ducar, C., & Schocket, D. H. (2018). Machine translation and the L2 classroom: Pedagogical solutions for making peace with Google translate. *Foreign Language Annals*, 51(4), 779–795. <https://doi.org/10.1111/flan.12366>
- Edwards, R., & Holland, J. (2013). *What is qualitative interviewing?* A&C Black.
- Hill, R. (1982). *A dictionary of false friends*. MacMillan.
- Hinkel, E. (2006). Current perspectives on teaching the four skills. *TESOL Quarterly*, 40(1), 109–131.
- Jones, A., & Bissell, C. (2011). The social construction of educational technology through the use of authentic software tools. *Research in Learning Technology*, 19(3). <https://doi.org/10.3402/rlt.v19i3.17116>
- Kline, R., & Pinch, T. (1996). Users as agents of technological change: The social construction of the automobile in the rural United States. *Technology and Culture*, 37(4), 763–795. <https://doi.org/10.2307/3107097>
- Lee, S.-M. (2020). The impact of using machine translation on EFL students' writing. *Computer Assisted Language Learning*, 33(3), 157–175. <https://doi.org/10.1080/09588221.2018.1553186>
- Machida, S. (2011). Translation in Teaching a Foreign (Second) Language: A Methodological Perspective. *Journal of Language Teaching & Research*, 2(4).
- Niño, A. (2009). Machine translation in foreign language learning: Language learners' and tutors' perceptions of its advantages and disadvantages. *ReCALL*, 21(2), 241–258. <https://doi.org/10.1017/S0958344009000172>
- Nishino, T. (2008). Communicative language teaching: An exploratory survey. *JALT Journal*, 30(1), 27. <https://doi.org/10.37546/JALTJJ30.1-2>
- Nivre, J., Abrams, M., Agić, Ž., Ahrenberg, L., Aleksandravičiūtė, G., Antonsen, L., Aplonova, K., Aranzabe, M. J., Arutie, G., Asahara, M., Ateyah, L., Attia, M., Atutxa, A., Augustinus, L., Badmaeva, E., Ballesteros, M., Banerjee, E., Bank, S., Barbu Mititelu, V., ... Zhu, H. (2019). Universal dependencies 2.4. <http://hdl.handle.net/11234/1-2988>
- Pym, A., Malmkjær, K., & Gutiérrez-Colon, M. (2013). *Translation and language learning: The role of translation in the teaching of languages in the European Union*. Publications Office of the European Union. <https://doi.org/10.2782/13783>
- R: A Language Environment for Statistical Computing. (2020). R Foundation for Statistical Computing.
- Silverman, D. (2013). *Doing qualitative research: A practical handbook*. SAGE Publications limited.
- Slobin, D. I. (1996). From “thought and language” to “thinking for speaking.” *Rethinking Linguistic Relativity*, 17, 70–96.
- Stapleton, P., & Kin, B. L. K. (2019). Assessing the accuracy and teachers' impressions of Google Translate: A study of primary L2 writers in Hong Kong. *English for Specific Purposes*, 56, 18–34. <https://doi.org/10.1016/j.esp.2019.07.001>

- Tsai, S.-C. (2019). Using Google translate in EFL drafts: A preliminary investigation. *Computer Assisted Language Learning*, 32(5–6), 510–526. <https://doi.org/10.1080/09588221.2018.1527361>
- Widdowson, H. (2003). *Defining issues in English language teaching*. Oxford University Press.
- Wijffels, J. (2020). udpipe: Tokenization, parts of speech tagging, lemmatization and dependency parsing with the “UDPipe” “NLP” toolkit [Manual]. <https://CRAN.R-project.org/package=udpipe>