The Abilities of Online Machine Translation and Resulting Implications for English for Academic Purposes

The Abilities of Online Machine Translation and Resulting Implications for English for Academic Purposes
オンライン機械翻訳ツールの能力と English for Academic Purposes 教育への影響

Klaus MUNDT* and Mike GROVES†
クラウス・ムント、マイク・グルヴズ

Abstract

Online Machine Translation (OMT) tools such as Google Translate have improved dramatically in recent years. Given the ease of access and the quality of the translation output, it seems reasonable to assume MT is used by students who are taught and assessed in the medium of English, but for whom English is not the first language; for instance to help them read or produce texts in English. This talk summarised observations on the development of online Machine Translation quality since 2014 and considered student use of MT in terms of the implications it may have for English for Academic Purposes (EAP) in Higher Education.

Google 翻訳のようなオンライン機械翻訳 (Machine Translation, MT) ツールは、近年飛躍的な進歩を遂げている。MT は、手軽に質のよい翻訳を得られるため、英語による授業を受講する英語を母国語としない学生が、英文の読解や作成の際に利用していると思われる。本講演は、2014 年以降のオンライン機械翻訳ツールの発展と進歩を総括するとともに、学生の MT 利用が大学教育の English for Academic Purposes (EAP) にもたらす影響について考察した。

1. Introduction

Although a relatively new technology, online machine translation (OMT) is already clearly having an impact on the teaching and assessment of English for Academic Purposes (EAP) (Clifford et al., 2013; Jolley & Maimone, 2015; Mundt & Groves, 2016; Tsai, 2019). In addition, there is a growing body of literature around the nature of this impact, of which Lee (2021) provides an excellent overview.

Given that it is virtually free of charge, instant and available to anyone with a web connection, it would seem fair to argue that it will become ever more embedded in the day to day practices of university students who are studying, or aspiring to study, in an environment whose language is not their first. Previous assumptions over exactly what and how these students need to learn are being challenged with this new technology (Kol, Schcolnik, & Spector-Cohen, 2018). From this, questions arise in a number of areas. These include but are not limited to the following.

* University of Nottingham
† University of Surrey

Do students need to spend time and effort becoming fluent in the grammatical systems of a foreign language, if the same effect can be achieved by the use of technology? Are students obliged to develop the receptive skills necessary to understand textbooks and lectures in a second language, if the technology is able to obviate this need? Can work submitted by students, but written with heavy reliance on OMT be said to be their own? What are the views of the non-EAP stake holders, such as academic staff in these debates?

This paper will address these questions by giving a summary of some of the research we have carried out over the last 7 years and suggesting some further areas for debate in this area.

2. The quality of Online Machine Translation

The first work we conducted on OMT in EAP involved looking at the results of translating students' writing in their first language through Google Translate into English (Groves & Mundt, 2015). We asked foundation (pre-undergraduate) students to write a simple argumentative essay in their first language, which was either Chinese or Malay. We then translated this through Google and counted the errors. At this point, it is important to note that we were not examining any other aspects of essay quality, for example argumentation, linking or organisation. The errors were counted and then compared with a paper by Müller (2015), who had counted errors in IELTS essays of various levels. We concluded that, from an accuracy point of view, Google was able to write at a level of around 6.0–6.5 IELTS. This meets the threshold for acceptance at many Anglophone universities.

At this point the distinction between absolute and relative quality became apparent. It was very clear that the language translated by Google was not written by an expert user of the language. However, at the level of 6.5 in IELTS, it was of higher quality that the student was able, in many cases, to produce unassisted (Mundt & Groves, 2016).

After a few years, we had noticed a radical improvement in the quality of Google's output. This could be explained by the implementation of neural network technology by Google (Turovsky, 2016). In 2018 we decided to rerun the original essays from 2014. The results were very noticeable, as can be seen in the examples in Table 1, where very few grammatical issues (underlined in Table 1) remain in the 2018 data.

Table 1 – Google Translate output changes over 4 years

2014	2018
In my opinion , the examination is very important to evaluate the individual but its relevance to today's society, there are limitations. Early age , the students were taught to learn in order to achieve excellent results by their parents , but in reality, not so easy to accept. Intelligence or excellence of a person can not be assessed through exams. Activities such as off-site is very dominant and practical lessons for the students to learn something . Such evaluation also plays an important role in testing and evaluation of students' skills . Activities that stimulate the minds of students can teach them to be critical and creative thinking . The only school -based learning in the classroom and on exams memory , will only make the learning process is not very efficient and effective	In my opinion, the exam is very important to assess a person but its relevance to society now has its limit. Since childhood, students have been taught to learn to achieve brilliant decisions by their parents, but the reality is not so easy to accept. The skill or excellence of a person can not be evaluated by examination only. Activities such as out-of-school and practical sites are very dominant for students to learn something. Such assessments also play an important role in testing and assessing the skills of students. Activities that stimulate the minds of the students can teach them to be critical and creative. Schools that are based only on classroom learning and memory-based examinations will only make the learning process less efficient and effective.

This led us to look at some higher-level texts, and we found published abstracts from languages including Korean and Indonesian. The OMT translated abstracts showed very few grammatical or coherence issues. They were not totally error free, but were clearly comprehensible, and in some ways showed impressive alignment with norms of anglosphere academic writing, for example in the density of nominalisations. That said, it should be noted that abstracts are generally very normed in their construction, and we did not look at texts with more complex or loaded terminology. An example of a part of an abstract translated from Indonesian follows to illustrate the quality of the OMT output:

> Second, this translation quality assessment model is well suited to assessing the quality of translation in the context of research and teaching professional translations. Third, this translation quality assessment model provides an opportunity for rater to provide translation assessment in various units, both at micro and macro level.

In the meantime, we also decided to investigate the ability of Google Translate to align itself to the norms and expectations of academic writing, beyond the level of grammatical accuracy. We approached this through the idea of Metadiscourse, as expounded by Hyland and Tse (2004). We took

a number of published German texts in the area of the humanities and translated them into English, again using Google Translate. We then counted two metadiscoursal features: hedges and boosters. Hedges are where a writer exercises caution in the expression of their beliefs. Examples include phrases such as “it seems” or “it could be argued that”. Boosters do the opposite. They give strength to an author’s beliefs. Examples include “definitely” or “certainly”.

After counting, we compared our finding with published studies analysing human writers. The results are summarised in Table 2. The exact numbers do not carry the main finding here, since all papers had a slightly different counting methodology. What is important is that in the studies concerned with human writers, there are more hedges than boosters (H>B), while the OMT translated texts contain more boosters than hedges (B>H). Therefore, while OMT was radically improving, it was not able to adapt to discourse norms beyond the level of the sentence.

Table 2 – Hedges and Boosters. A comparison between human and GT translations from German

	Hedges per 1000 Words	**Boost per 1000 words**	**Summary**
Hyland (2005) – Philosophy	12.3	8.3	H>B
Hyland and Tse (2004) – Applied Linguistics	11.1	3.8	H>B
Mur-Dueñas (2011)	20	6.4	H>B
Our study	5.4	9.1	B>H

Overall then, it is clear that OMT has developed greatly in the last 7 years. The system is able to write with impressive grammatical accuracy, far in advance of the threshold level of university study. However, other, deeper, aspects of EAP are not dealt with well by this system. Put simply, it can handle the E but not the AP.

3. Technological Context

This also needs two pieces of contextualization. Firstly, technology surrounding the development of OMT is developing equally quickly. For example, MS Teams is rapidly developing its ability to caption live conversations, and AI systems such as Otter.ai are able to provide impressive transcription. It is not hard to envision a system that combines these technologies, to provide live, translated subtitles in lectures. One objection could be raised that automatic transcription and translation are far from perfect in their accuracy, and such a system could easily bring poorly translated information to the students. While this is the case, it returns the argument to the second piece of contextualization—which is the idea of absolute versus relative quality (Groves and Mundt, 2021). It is, of course, true that an online system could lead to misunderstanding. However, it must be borne in mind that for the vast majority of students who are studying in a foreign language, such misunderstandings are an inevitable part of their study until they gain a level of expertise which approaches a native speaker—and this is not something the vast majority of students achieve.

4. Theoretical Context

It is worth examining the context beyond the area of EAP, and even beyond Higher Education. Technology has been influencing and disrupting various social practices, perhaps since the invention of the needle. This was popularised, in the 1930s and 40s by the economist Joseph Schumpeter as "Creative Destruction". This is the way in which one technology will displace a previous one, and is, according to the theory an ingrained part of capitalism. Examples abound, and there are sometimes obvious and less obvious winners and losers in the process. For example, when music moved from vinyl to CD in the 1980s and 1990s, it did not just affect the music industry itself. It also affected chemical firms whose business model depended on the manufacture of vinyl (Pettinger, 2018).

More recently, this has been theorised as "Disruptive Innovation". Daneels (2004) explains that this process is rarely instant. At first the new technologies have little impact, and this is often down to poor performance. However, over time, the technology improves to the point where it moves into the mainstream. Millar, Lockett, and Ladd (2018: 254) define the process as a "change that makes previous products, services and/or processes ineffective". It is not hard to envisage OMT on this trajectory—moving from ineffective marginal technology into the mainstream.

A second relevant theory is known as the Technology Acceptance Model (TAM) (Davis, 1989), which seeks to describe how technology moves into mainstream acceptance. At the risk of oversimplification, the TAM asks two questions. The first of these is whether the new technology is easy to use. OMT is clearly easy to use—notwithstanding questions around OMT literacy, discussed below. Written communication of all types can be translated instantly with a device almost everyone owns. The translations are not always perfect, but they convey meaning well.

The second question is whether the technology can help the user perform their role better. The answer to this question is more nuanced. If the "role" of the EAP student is to simply learn English, then the answer to the question is 'maybe'. Used well, OMT can clearly be used as a useful supplement for students' self-study. However, it can also be used to help students avoid the need to learn complex English, especially for assessment. However, it is a highly questionable, and possibly colonial, assumption that EAP students' ultimate goal is to learn English. There is nothing intrinsic in English that makes it a worthy goal in itself. English is the academic Lingua Franca because a Lingua Franca is needed if there is to be a global community of scholarship, and English fulfils that role for reasons beyond the scope of this paper. For the vast majority of EAP students, English is a means to an end, not the end in itself. If we see the students' goal as participation in a global academic community, then the answer to the second question must be an unequivocal 'yes'. Should the technology develop to a point where face to face communication through OMT can approach the ease of written communication, it could lead to a situation where international academic communication ceases to take place in English. Scholars would simply communicate in their own language and rely on computer systems for understanding.

5. Online Translation in Higher Education

The acceptance of technology in Higher Education (HE) is not only decided by the users but also by university policy. Since OMT has only relatively recently become sophisticated enough to be of use to the student user, we conducted a study to investigate the attitude towards student use of OMT among academic staff (see Groves & Mundt, 2021 for detailed findings). The study was conducted at two UK Russell Group universities and collected data by means of 13 interviews. The participants ranged from lecturers to full professors from a range of disciplines and also included members of academic services (i.e., non-teaching/research staff) involved in quality assurance and assessment and academic integrity policy making. Staff from Modern Language departments and EAP units were not included in this study, since it seems safe to assume that they would be mainly concerned about students' language acquisition. This would certainly be worth a separate investigation. However, we wanted to obtain a picture of the wider academic community, for whom the language aspect is not usually the focus of attention. The following main themes emerged from the data:

- Policy
- Control
- Academic Integrity/Writing
- The wider academic community/Employability
- Transition

The participants were in agreement that there was, at the time of data collection, no policy governing the use of OMT by students. Equally, there was the recognition that it may not be possible to control the use of OMT.

Regarding academic integrity, the views were more nuanced. First of all, the participants did not voice greater reservations concerning the use of OMT as supporting tool for reading. In other words, there seemed to be little concern about a student whose first language is not English to use OMT to help them engage with the literature. This was different for the production of assessed writing, such as course work or essay assignments. The participants seemed more at ease with the use of OMT at the micro-level. That is to say, there seemed to be fewer concerns about a student writer using OMT as substitute for a dictionary or as writing aid to produce shorter stretches of text, e.g., by writing a sentence in the first language and translating it into English.

There was more concern about a student potentially writing an entire paper in their first language and using OMT to convert it into the final English draft for assessment. Several factors were at play in this regard. One recurrently mentioned concern was the status of English and the university brand. A number of participants pointed out that a degree from a UK university implies that the graduate has a working command of English—and external stakeholders, such as employers of those graduates, will rely on that to be the case. The other concern was that reliance on OMT could prevent, for instance, research students from participating in events such as presenting at conferences.

The data suggested that the participants envisaged that the use of OMT would decline as the

student progresses through their studies—so OMT was perceived by some as a viable learning tool at the early stages of their learning process, but less so at the later stages. However, it was not specified where the cut off might be in this transition.

Importantly, only one participant considered that the reliance on OMT to produce assessed written work constituted academic misconduct. This was, in fact, contingent on the market expectations that UK graduates speak English well. The other participants did not share this view, stating that the use of OMT, despite all reservations, did not constitute academic misconduct.

In summary, then, the participants were not generally opposed to OMT, but also not ready for it to take over. Since there seemed to be cautious acceptance that OMT is here to stay, that it will improve, and that students are likely to use it, it seems prudent to consider it as a tool, the responsible use of which might need to be taught and thus be viewed as part of academic literacies, specifically digital literacies for HE.

5.1 Online Translation as part of digital literacies

The JISC Framework for Digital Literacies (2015) is quite comprehensive, but does at present not include OMT. This might be because it originates from a time when OMT was not really quite yet a 'serious player' in HE.

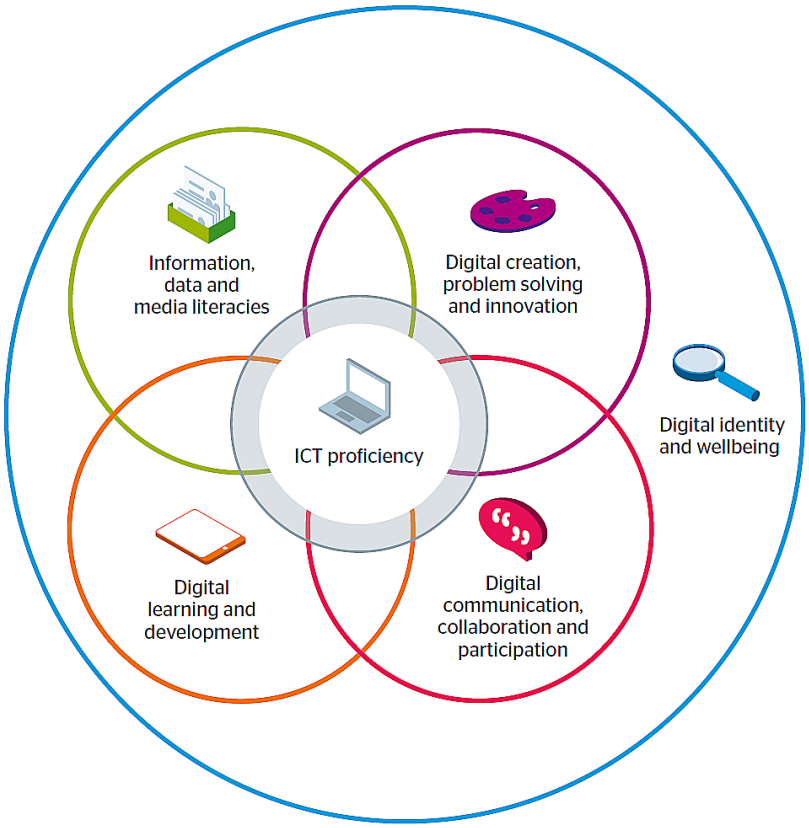

Figure 1 – JISC Framework for Digital Literacies

Hence, in all its areas, OMT simply remains a blind spot despite the fact that students use it. Perhaps it could fall into the areas of digital problem solving or digital communication and participation, which are at present more viewed as technical skills, (e.g., how to work an app, how to use cloud spaces for collaborative work, etc.) as opposed to facilitating communication (i.e., language processing (reading & writing)). Either way, it would seem prudent to recognise the capabilities of OMT for future iterations of this framework or others like it.

Bowker and Ciro (2019), in recognition of the continuously increasing quality and use of OMT, presented a number of suggestions on how to write in a machine translation friendly way, so as to facilitate the accurate and clear rendition of texts via machine translation:

- Use short sentences
- Use the active voice rather than the passive voice
- Avoid long noun strings & modifier stacks
- Use relative pronouns such as 'that' & 'which'
- Avoid wordiness
- Use nouns instead of personal pronouns
- Use terminology consistently
- Choose unambiguous words
- Avoid abbreviated forms
- Avoid idiomatic expressions, humour & cultural references

While these suggestions are indeed likely to raise the quality of OMT output, they also raise a number of questions. Since these guidelines are rather restrictive, there may be resistance from academic writers—because writing in a machine translation friendly way may require them to change ingrained writing habits and undermine or contradict personal preferences and stylistic preferences.

Another important point to consider is that Bowker & Ciro's machine translation literacy does not take the EAP community into account at all (Mundt & Groves, 2019). Their focus is on writing for publication—this also means they never consider elements of academic integrity, where arguably a translated text may be considered not of the student's own making.

Also, in most EAP contexts, the student writer is likely to be less competent in academic discourse and language than the teacher reader. In other words, it may be difficult for the student user to strategically follow the guidelines. It may also undermine the development of their own writer's voice and style and may result in a new register of machine-generated English. At the same time, this illustrates that OMT is currently not poised to replace language acquisition.

5.2 Online Translation in EAP

In light of the rapid improvements of OMT output quality, and in light of its user-friendliness, we are no longer asking ourselves if OMT should be integrated into the EAP curriculum—especially with

a view to the future of EAP. We are convinced it will stay with us and it is bound to continue to improve. Given the pressures that our students face, it seems only reasonable to assume that they will continue to use OMT, just like dictionaries once became an integral part of the 'survival kits' of L2 learners. Thus, we are now thinking on the lines of how this integration of the effective and responsible use of OMT into the curriculum could occur, and what that would mean for EAP, its students and its teachers.

For effective use, students need to understand how their L1 and Academic English relate through OMT mediation—so they will still require language awareness (at least declarative knowledge) of discourse preferences, e.g., metadiscoursal features such as hedging, boosting, use of sign-posts etc. More concretely, the students could experiment, while the EAP professional feeds back in the classroom or tutorials—this would then focus on instances of lack of clarity, etc.—like in any kind of formative feedback on writing, where the students then also develop a better awareness of how the machine can help them (if at all), and make better informed decisions what they want to use it for (if anything). The subtler the students' understanding of the technology and the language, the better the writing will be, and there remains viable scope for the technology to be phased out as a support mechanism, possibly even by the students themselves.

Some key aspects of current EAP teaching would remain the same. For instance, aspects such as avoiding overly complex sentences or avoiding ambiguity are already central tenets of academic writing literacy. This would remain so, but with a heightened sense of working with technology.

Control over their writing would arguably move more to the student—i.e., through making them more literate, we can reduce over-reliance on and introduce the judicious use of technology, where own writing skills and technological facilitation blend into a discrete skill that the students have increasing control of.

It is in this light that we argue that OMT does not replace the need for EAP. We further argue that English language proficiency should still be assessed. This is due to the importance of incidental learning and also controlled assessment in the academy; and not least due to accountability to external stake holders such as visa-issuing entities. One limitation is that OMT literacy improves the ability to function with limited language ability, but only with texts. It is a tool—a facilitator, not a replacement for linguistic and inter-cultural competence.

We suggest that EAP faces a choice. We can either ignore OMT or we can engage with it in a transparent and constructive manner. If we ignore it, urgent questions will remain about student use of OMT. We would run this risk of inconsistency by, for instance, leaving decisions about acceptability to the arbitrary judgement of a course coordinator. There is also the risk of passing this issue on the academic misconduct officers, who, in the absence of policy, will have no answers either. This does not seem a particularly professional way of handling the issue. EAP should proactively participate in finding answers for the difficult questions it faces.

If we decide to engage with it, EAP has the opportunity of actively participating in important university policy making. Transparent and consistent engagement would also allow us to be consistent

and fair in what we do. However, this may entail some changes to established practice. Apart from lesson materials incorporating elements of OMT literacy, assessment would have to be adapted. One way of doing this for the assessment of writing abilities might be that for non-controlled work (e.g., essays) assessment criteria could place 'a focus on discourse […] embedded in social practices, disciplinary epistemologies' (Hyland 2018: 390).

Language proficiency could then be assessed in a more controlled environment. We might witness a re-emergence of the timed essay. But there are also alternatives. For instance, text editing could serve as an assessment opportunity. Another option might be to ask for written commentary on text (the students' own or other texts). The decision on how to assess language proficiency will have to be made locally. These need to be informed decision that also comply with university quality assurance and, ideally, the demands of external stakeholders. Essentially, then, we suggest that EAP can guide students in the judicious and responsible use of OMT as a transitional tool and learning facilitator—not a substitute for learning and communicating.

6. Conclusion

We suggest that it is evident that OMT has established itself as a tool that students will use when operating in a language that is not their first. The degree to which they will use it will differ, as will the specific purposes of use. We further suggest that EAP can take a leading role in determining useful ways of working with the technology. There is little point in vilifying it or arguing for students being barred from using it—not least since there currently seems no way of reliably controlling the use of OMT.

Of course, at present, there seem to be more questions than answer, and we would like to close by leaving the reader with a further few that may be useful to consider in the debate about how to handle OMT in the context of EAP or EMI (English as Medium of Instruction):

- If a student wants to be, e.g., an academic engineer what skills do they need?
- Is the dominance of English in academia just?
- Does English have any intrinsic value that makes it the academic Lingua Franca?
- In times of 'global English', by whose standard do we measure 'English proficiency'? (cf. Widodo & Fang, 2019)
- How sure are we that Artificial Intelligence can never adapt metadiscoursal functions in OMT?
- Is there more to EAP than the language?
- Does EAP become CoMLAP? —Computer Mediated Language for Academic Purposes?

References

Bowker, L., & Ciro, J. B. (2019). *Machine translation and global research*. Bingley: Emerald Publishing.

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*, *10*, 108–121. https://doi.org/10.7203/attic.10.2228

Daneels, E. (2004). Disruptive technology reconsidered: A critique and research agenda. *Journal of Product Innovation Management*, *21*, 246–258. https://onlinelibrary.wiley.com/doi/epdf/10.1111/j.0737-6782.2004.00076.x

Davis, F. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly*, *13*(3), 319–339. https://doi.org/10.2307/249008

Groves, M., & Mundt, K. (2015). Friend or foe? Google Translate in language for academic purposes. *English for Specific Purposes*, *37*, 112–121. https://doi.org/10.1016/j.esp.2014.09.001

Groves, M., & Mundt, K. (2021). A ghostwriter in the machine? Attitudes of academic staff towards machine translation use in internationalised Higher Education. *Journal of English for Academic Purposes*, *50*, 100957. https://doi.org/10.1016/j.jeap.2021.100957

Hyland, K. (2005). *Metadiscourse: Exploring interaction in writing.* London: Continuum.

Hyland, K. (2018). Sympathy for the devil: A defence of EAP. *Language Teaching*, *51*(3), 383–399. https://doi.org/10.1017/S0261444818000101

Hyland, K., & Tse, P. (2004). Metadiscourse in academic writing: A reappraisal. *Applied Linguistics*, *25*, 156–177. https://doi.org/10.1093/applin/25.2.156

JISC. (2015). Developing students' digital literacy. Retrieved 15 October 2019 from https://www.jisc.ac.uk/guides/developing-students-digital-literacy

Jolley, J. R., & Maimone, L. (2015). Free online machine translation: Use and perceptions by Spanish students and instructors. In A. J. Moeller (Ed.), *Learn language, explore cultures, transform lives* (pp. 181–200). WI: Robert M. Terry.

Kol, S., Schcolnik, M., & Spector-Cohen, E. (2018). Google Translate in academic writing courses? *The EuroCALL Review*, *26*(2), 50–57. https://doi.org/10.4995/eurocall.2018.10140

Lee, S.-M. (2021). The effectiveness of machine translation in foreign language education: A systematic review and meta-analysis. *Computer Assisted Language Learning*. https://doi.org/10.1080/09588221.2021.1901745

Millar, C., Lockett, M., & Ladd, T. (2018). Disruption: Technology, innovation and society. *Technological Forecasting and Social Change*, *129*, 254–260 https://doi.org/10.1016/j.techfore.2017.10.020

Müller, A. (2015). The differences in error rate and type between IELTS writing bands and their impact on academic workload. *Higher Education Research & Development*, *34*(6), 1207–1219. https://doi.org/10.1080/07294360.2015.1024627

Mundt, K., & Groves, M. (2016). A double-edged sword: The merits and the policy implications of Google Translate in higher education. *European Journal of Higher Education*, *6*(4), 387–401. https://doi.org/10.1080/21568235.2016.1172248

Mur-Dueñas, P. (2011). An intercultural analysis of metadiscourse features in research articles written in English and in Spanish. *Journal of Pragmatics*, *43*(12), 3068–3079. https://doi.org/10.1016/j.pragma.2011.05.002

Pettinger, T. (2018). *Creative destruction*. Economics Help. Retrieved 12 January 2022 from https://www.economicshelp.org/blog/20255/economics/creative-destruction/

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

Turovsky, B. (2016). Found in translation: More accurate, fluent sentences in Google Translate. *The Keyword*. Retrieved 12 January 2022 from https://www.blog.google/products/translate/found-translation-more-accurate-fluent-sentences-google-translate/

Widodo, H. P., & Fang, F. (2019). Global Englishes-oriented English language education. In F. Fang & H. P. Widodo (Eds.), *Critical perspectives on Global Englishes in Asia* (pp. 194–200). Bristol: Multilingual Matters.