Machine translations revisited: issues and treatment protocol

Harry Harris
Hakuoh University

Machine translation (MT) is a pedagogical issue in second language (SL) writing because student use means less engagement with the target language and diminished chances to learn it. It is also a quality issue because current MT technology is imperfect. This paper examines these issues, arguing that the current research-backed process writing methodology does not generally support student submissions of MTs for writing assignments. It also demonstrates several language problems resulting from MTs. Although MT can be valuable for business, government, and personal purposes, SL writing teachers should discourage student use of the technology.

機械翻訳（MT）は第2言語（SL）ライティング学習における教育上の問題となっている。その使用により、学習者が目標言語（TL）に関わることがより少なくなり、その言語を学ぶ機会を減少させるからである。また、現在のMT技術は不完全であり、その質も問題である。本論ではこれらの問題を考察し、現在の研究に裏づけされたプロセス・ライティング方法論は一般的に、学習者のMTを使ったライティング課題提出を支持しないと論ずる。また、MTにより生ずる数々の問題も実証する。MTはビジネス、行政や個人的な目的には大変役立つかもしれないが、SLライティング指導者は、学習者にMT使用を止めさせるようにすべきである。

Several of my EFL students had just come to my office with a belated group-effort assignment which, after skimming through, I found incomprehensible. What caught my eye immediately were syntactic and semantic issues that left me clueless. Some of the vocabulary was contextually peculiar, and there were pronouns with obscure referents. I knew the challenge English posed for these students, and I could usually identify their typical errors, but these textual problems frustrated me. I began to wonder whether I was looking at a machine translation (MT), so I asked them about this. Abashed, they replied *chotto* (a little). I reminded them of our talk at the beginning of the semester about plagiarism, homework borrowing, and the pedagogical and poor-quality issues of MTs. They agreed to rewrite their assignment and left.

MT-suspect writing assignments such as this often crop up in our second language (SL) classes, a situation requiring vigilance and appropriate treatment. We understand our students are busy and that the temptation to write their assignments in L1 and do an MT can be strong.[1] However, there are pedagogical issues signaling that MT use can rob students of opportunities to use the target language meaningfully and quality issues that can frustrate and cause communication to collapse. Moreover, many teachers are convinced that MT submissions are a form of dishonesty by students disrespectful of the assignment and are unfair to those who laboriously craft their writing (Zemach, 2008). Also, as Wheeler (2009) argues about plagiarism, this practice of cheating cannot be solely attributed to inherent

---

1 MT refers to any free or paid, online or offline automatic translation service to which students have easy access. Common services include Yahoo! Babel Fish and WorldLingo.

cultural values demanding teacher tolerance. Finally, we may feel annoyed when forced to take time determining whether unusual text in student assignments is due to MT or English-language problems. This is not always easy to determine, especially with basic English students. Discerning MT text from natural errors requires a careful response married with respect (Silva, 1997), transparency (Zemach, 2008), and guidance (Currie, 1998) despite inconsistent definitions of academic dishonesty (Pincus & Schmelkin, 2003) and the uncertainties inherent in acting on this Pandora's Box (Currie, 1998).

To better understand these issues and our own need to respond resolutely to student MT submissions, this paper provides a brief reminder of the process approach to current SL writing pedagogy, a tentative analysis of MT inadequacy, and a treatment protocol for this persistent problem.

Pedagogical issues

Process writing, rather than product, should be the focus of SL writing education (Zamel, 1976). Traditional product writing methodology often consists of single-draft assignments assessed on their touchstone-like adherence to model essays and to instructor-held ideas about writing quality. These assignments are collected and graded, and then students go on to the next assignment. Process writing, however, focuses on the cognitive and social nature of writing. It is a recursive process where writers periodically reconsider and revise earlier text for further development of ideas and critical thinking skills. In this approach, students require formative feedback and invention strategy instruction in order to maximize learning (Matsuda, 2003).

Although not for process writing, MT use can be perfectly valid in certain contexts. For example, multinational companies have produced MTs of professional documents since the 1950s and 60s (Hutchins, 2007). In reference to patent translations, MTs can also be much cheaper than human translators (Vitek, 2000). Additionally, the quick and steady production of MTs can prove vital in fields such as meteorology, as demonstrated by the huge number of daily weather bulletins translated by the METEO system in Canada (Napier, 2002).

However, despite the growth of the Internet and the increasing user need to obtain information from foreign language sources, MT should not be an SL shortcut. This does not mean, of course, that our students should not obtain information from first language sources that will help them develop writing topics. Nor does it mean that SL writing instructors should not, for example, design editing or other activities around MTs (Niño, 2009). All these can be fruitful endeavors. It does mean, though, that without learning to manipulate its distinctive writing conventions, SL students will not learn to write effective English on their own.

Quality issues

Despite some MT support in non-SL process writing education contexts and the presence of MT-supportive organizations in Asia (AAMT), Europe (EUROTRA), and the Americas (AMTA), many experts realize current MT limitations. Simply put, MT does not work well unless there is sufficient world knowledge (knowledge of the relationship between things) and controlled language, as in this example of a French-to-English MT (Napier, 2002): The MT turns *Les soldats sont dans le café* into *The soldiers are in the coffee,* mistaking the beverage for the shop, effectively placing people in it. As Melby (1995, p. 4) writes, "computers are very likely to produce atrocious results [unless the text treats] a very narrow topic in a rather dry and monotonous style."

At this stage in the development of MT technology, the consensus is that MT requires an enormous amount of human intervention to maintain quality control. As Vitek (2000, 2001) notes, since a machine does not understand concepts of accuracy or meaning and lacks agency (ability to make a choice exercising one's will), accurate translations may only be possible if the input range is extremely limited, with the input, translation process, and output subject to a strict and *professional* [emphasis added] human control. Without this control, MT cannot deal with real vocabulary surprises (Melby, 1995).

Natural human language is just too complex for accurate unassisted contemporary MT. Although we may be able to program computers with syntax-coping capability, MT theorists and those working in the fields of neural networks

and artificial intelligence are still unable to give computers knowledge of meaning (Napier, 2002). As Budiansky (1998, para. 22) reminds, "Language is full of ambiguity and multiple meanings that a correct reading of syntax goes only a short way toward sorting out [because] computers don't have any common sense."

With MTs, however, syntax and semantics are *both* problems (as is rhetorical style, an unaddressed issue here), especially with languages as distinct as English and Japanese. An increase in grammar complexity, lexeme ambiguity, as well as idiomatic and culture-specific language creates conditions ripe for MT meltdowns, as a brief comparison of Yahoo! Babel Fish (2008) and WorldLingo (2008) MTs in Table 1 shows. Here A refers to the original sentence, B to the MT into Japanese, and C to the back-translation (BT) into English.

Apart from a few minor differences, the Yahoo! Babel Fish and WorldLingo BTs are comparable. Also, there are no major differences with the Japanese MTs.

Of interest is how the Japanese MTs and BTs diverge from the original English sentences. The Japanese MTs of Sentence 1 provide a thoroughly unexpected change in meaning. Whereas the original sentence expresses the idea of Bob's love for Clara and her love for Bill, the Japanese MTs have changed this to Bob's love for Bill *and* Clara, or *Bob loves Bill and Clara*. Perhaps this has happened because both Clara and Bill follow Bob syntactically. The BTs are, of course, deplorable.

Both Bob and Bill are no longer people, with Bob becoming *the hob* and Bill *the building*. Although the Japanese translation does represent Bob as ボブ [*bobu*], the translators are not reading the diacritic (ﾞ) that converts the Japanese *ho* [ホ] into *bo* [ボ]. With the latter name, the Japanese translations have Bill as its Japanese homophone ビル [*biru*], meaning *building*, a change resulting, perhaps, from the fact that there is no distinction between upper and lower case script in Japanese.

Sentence 2 offers an example of a mistranslated English phrasal verb, *break down*, which has several different transitive and intransitive meanings (*Shorter Oxford English Dictionary*, 2003, p. 284). Though the idiom in the original English sentence is intransitive and contains the meaning *loss of self-control*, the Japanese translation 破壊 [*hakai*] is used transitively by Babel Fish and intransitively by WorldLingo and carries the English meaning *destroy*. Additionally, Babel Fish translates the word cried into Japanese as 叫んだ [*sakenda*], meaning shouted. WorldLingo's translation of this word comes out as 叫ぶ [*sakebu*], or *shout*. The BTs, of course, are barely intelligible.

With Sentence 3, the MTs fail to recognize that in English *June* can be the name of a person as well as a month. Both translators consistently render this proper noun as 6月 [*rokugatsu*], the month of the year, making the Japanese translations unacceptable. The BTs can only be interpreted as back-transformations of the month.

Though the above MT examples need further

Table 1. Comparison of Yahoo! Babel Fish and WorldLingo MTs

	Yahoo! Babel Fish	WorldLingo
1	A. Bob loves Clara, who loves Bill. B. ボブはビルを愛するクララを愛する。 C. The hob loves [kurara] which loves the building.	A. Bob loves Clara, who loves Bill. B. ボブはビルを愛するClaraを愛する. C. The hob loves Clara which loves the building.
2	A. Sam broke down and cried. B. サムは破壊し、叫んだ。 C. It destroyed sum, shouted.	A. Sam broke down and cried. B. 破壊され、叫ぶサム。 C. The sum which is destroyed, shouts.
3	A. June likes cake. B. 6月はケーキを好む。 C. June likes the cake	A. June likes cake. B. 6月はケーキを好む。 C. June likes the cake.

corroboration with other MT platforms for performance comparison, at least with WorldLingo and Babel Fish, there can be egregious problems. The Japanese translations don't always match the syntactic intentions of the original English sentences, and there are semantic issues as seen in the mistranslations. The BT sentences inconsistently range from completely acceptable (though still unable to match the Japanese in intended meaning), to odd, to utterly incomprehensible because of script differences and other reasons. It is clear that MTs can be problematic, at least with Japanese.

Treatment protocol

As with plagiarism and assignment borrowing, teachers should, as Gerdeman (2000) indicates, inform students through a variety of means (including course-initial announcements and syllabi) that unless there is a specific purpose for them, MTs are unacceptable and will have a detrimental effect on the learning process. This first point warns students and provides justification for future grade assessment; the second educates them or at least provides a message that may one day be meaningful. Importantly, students should know that submitting writing assignments containing MTs will negatively affect their grade.

An effective awareness-raising activity about MT problems is to have students correct a teacher-provided English-to-L1 MT (Harris, 2009). After spending some time on this exercise, students will soon realize that MTs are not a viable alternative to their own work (because the teacher will know they did not actually write anything) and will hopefully not attempt to use them later.

Despite these deterrents, there will probably be students whose minds wandered when course-initial explanations were made or who just did not find the above class activity credible. Thus, at some time during the course, MT-suspect assignments may appear. Though it is indeed time-consuming to deal with this issue, we cannot ignore it. Teachers who receive MT suspects should repeat their MT class warning and inform students of the comprehension problems with their work. If students do not respond meaningfully, the teacher can then point out specific text

problems and ask for an explanation. If the students cannot, ask them to rework the assignment until it is satisfactory, which in my experience usually results in a non-MT submission. Remind these future teachers and business people that MT is far from perfect and that MT renditions of any written work can, unless competently edited, present problems for those who read them, with consequences that can range from reader miscomprehension and withdrawal to customer exasperation and desertion.

Conclusion

MT, whatever the consequences, is something technological developments have made available. There are circumstances in which MT plays an important role, saving money and time, in the endeavors of business, government, and other entities. However, the possibility of imperfect translations requires careful human attention and intervention. MT access allows us to obtain international information and communicate in some form with others who speak different languages. Those who use this technology, especially our SL students, need to be aware of its current performance imperfections. Above all, our students need to know that when they resort to an MT service in order to avoid doing their own writing, they lose a valuable opportunity of learning how the language functions. They will remain isolated both from the language and from those with whom they would communicate. We should take them to task on this.

Acknowledgements

References

Budiansky, S. (1998). Lost in translation. *The Atlantic Online*. Retrieved February 14, 2009, from <www.theatlantic.com/issues/98dec/computer.htm>.

Currie, P. (1998). Staying out of trouble: Apparent plagiarism and academic survival. *Journal of Second Language Writing, 7*(1), 1-18.

Gerdeman, R. D. (2000). Academic dishonesty and the community college. (ERIC Document Reproduction Service No. ED447840).

Harris, H. (2009). The pedagogical frustrations of machine translations: Vain areas without shape. *SLW & CALL Perspectives: An InterSection Newsletter of TESOL, 1*(1). Retrieved Feb. 7, 2010, from <www.tesol.org//s_tesol/sec_issue.asp?nid =3124 &iid =12242&sid=1>.

Hutchins, J. (2007). Machine translation in Europe and North America: Brief account of current status and future prospects. *Japio 2007 Year Book*. Tokyo: Japan Patent Information Organization, 222-227. Retrieved February 7, 2010, from <japio.or.jp/00yearbook/files/2007book/07_4_03.pdf>.

Matsuda, P. K. (2003). Second language writing in the twentieth century: A situated historical perspective. In B. Kroll (Ed.), *Exploring the dynamics of second language writing* (pp. 15-34). New York: Cambridge University Press.

Melby, A. K. (1995). Why can't a computer translate more like a person? *Fortune City*. Retrieved February 14, 2009, from <www.mt-archive.info/Melby-1995.pdf>.

Napier, M. (2002). The soldiers are in the coffee - An introduction to machine translation. *Cultivate Interactive*, *2*. Retrieved February 12, 2009, from <www. cultivate -int.org/issue2/mt/>.

Niño, A. (2009). Machine translation in foreign language learning: Language learners' and tutors' perceptions of its advantages and disadvantages. *ReCALL, 21*(2), 241-258.

Pincus, H. S. & Schmelkin, L. P. (2003). Faculty perceptions of academic dishonesty: A multidimensional scaling analysis. *The Journal of Higher Education, 72*(2), 196-209. Retrieved February 8, 2010, from <www.swarthmore.edu/NatSci/cpurrin1/plagiarism/docs/Pincus2003.pdf>.

Shorter Oxford English Dictionary. (2003). Fifth Edition. New York: Oxford University Press.

Silva, T. (1997). On the ethical treatment of ESL writers. *TESOL Quarterly, 31*(2), 359-63.

Vitek, S. V. (2000). Reflections of a human translator on machine translation or will MT become the Deus Ex Machina rendering humans obsolete in an age when Deus Est Machina? *Translation Journal, 4*(3). Retrieved February 12, 2009, from <accurapid.com/journal/13mt.htm>.

Vitek, S. V. (2001). Useful machine translations of Japanese patents have become a reality. *Translation Journal*, *6*(2). Retrieved February 14, 2009, from <accurapid. com/journal/20mt.htm>.

Wheeler, G. (2009). Plagiarism in the Japanese universities: Truly a cultural matter? *Journal of Second Language Writing, 18*(1), 17-29.

WorldLingo. (2008). Free Online Language Translator. WorldLingo Translations, LLC. Used for translation February 14, 2009, from <www.worldlingo.com/>.

Yahoo Babel Fish. (2008). Yahoo! Inc. Used for translation February 14, 2009, from <www.babelfish.yahoo.com/>.

Zamel, V. (1976). Teaching composition in the ESL classroom: What we can learn from research in the teaching of English. *TESOL Quarterly, 10*(1), 67-76.

Zemach, D. (2008). That's cheating! *Essential Teacher, 5*(3), 13-14.

Harry Harris, who has an MA in applied linguistics and one in Spanish linguistics, has taught ESL/EFL and Spanish at academic institutions in South America, Asia, and the United States. At present, he is a faculty member at Hakuoh University in Tochigi Prefecture, Japan. He also participates in teacher-training workshops, writes materials for publishing companies, and is a long-term English-language interviewer. A Hispanophile, Harry shares his summers with long-term friends in Granada, Spain.