ARTICLE

Student use and instructor beliefs: Machine translation in language education

Emily A. Hellmich, Berkeley Language Center

Kimberly Vinall, Berkeley Language Center

Abstract

The use of machine translation (MT) tools remains controversial among language instructors, with limited integration into classroom practices. While much of the existing research into MT and language education has explored instructor perceptions, less is known about how students actually use MT or how student use compares to instructor beliefs and expectations. In response to this gap, the current article explores how students use MT while writing and how this use compares to instructor perceptions via two studies: a computer-tracking study of how 49 second semester-level language learners (French, Spanish) use MT and a qualitative survey of 165 US-based second language educators' beliefs about MT. Findings highlight important areas of alignment (e.g., MT input at word level) and divergence (e.g., MT output analysis strategies) between student use and instructor perceptions as well as layered tensions in what mediates student use of MT tools. The article concludes with calls for more research on student use and an outline for how to approach MT tools in language education in ways that support existing student practices.

Keywords: *Machine Translation, Computer-tracking Study, Instructor Perceptions, Ecological Theoretical Approaches*

Language(s) Learned in This Study: *French, Spanish*

Introduction

Despite the rise in the ubiquity of machine translation (MT) tools as well as suggestions that MT can act as a supportive tool to language learning/teaching (Garcia & Pena, 2011; Lee, 2020; Tsai, 2019), the use of MT tools remains controversial among language instructors, with limited integration into classroom practices (Barr, 2013; Briggs, 2018; Hellmich & Vinall, 2021; Niño, 2009).

Much of the existing research into MT and language education has explored this disconnect via instructor and student beliefs (Case, 2015; Clifford et al., 2013; Jolley & Maimone, 2015; Niño, 2009). Less is known about how students actually use MT. Moreover, it is unclear how student use of MT compares with instructor perceptions, both of student use and of MT tools more broadly. From an ecological theoretical perspective, both how students use MT and how teachers perceive it mediate what happens when MT meets the language classroom.

To that end, the current article draws on two studies—a computer-tracking study of how second semester-level learners of French and Spanish use MT and a qualitative survey of US-based language educators' beliefs about MT—to put into conversation student use of MT and instructor perceptions. The overarching goals of this dialogue are to provide insights into how students use MT and how this use both aligns with and diverges from instructor expectations, so as to guide the development of future research and practice.

Background

Machine Translation and Instructor/Student Beliefs

Machine translation (MT) technologies—or the use of software to automatically translate text from one language to another (Qun & Xiaojun, 2015, p. 105)—has undergone rapid development in the past decade. The latest versions of MT rely on machine learning, rendering these versions faster, more efficient, and able to produce more accurate translations for certain languages (Kelleher, 2019; Lewis-Kraus, 2016; Poibeau, 2017; Wu et al., 2016). These changes have been documented by the MT industry (Lewis-Kraus, 2016; Wu et al., 2016) as well as language teaching/learning professionals (Briggs, 2018; Ducar & Schocket, 2018; Stapleton & Kin, 2019).

The potential implications of machine translation tools for language teaching and learning have been explored by technology-enhanced language learning researchers and practitioners. Extensive research has been done on what language instructors think about MT tools. A common center of this research base is the acceptability or ethicality of these tools. The available literature indicates mixed feelings on the part of instructors on this score. While one study reported that instructors find the use of MT to be a "non-serious" offense (Correa, 2011), others have found high rates of disapproval, including more than a third designating MT use as "cheating" (Clifford et al., 2013). A consistent trend in the literature connects the acceptability of MT with the parameters of its use. Text length (Case, 2015; Jolley & Maimone, 2015), assignment type (Clifford et al., 2013; Jolley & Maimone, 2015), and skill (Hellmich & Vinall, 2021) have all been found to impact instructor beliefs on acceptability of MT.

Another common topic examined in research on language instructor beliefs about MT is its role in the FL classroom. Overall, integration of MT into classroom practices has been reported to be limited (Barr, 2013; Briggs, 2018; Hellmich & Vinall, 2021; Niño, 2009). That said, many instructors believe that MT can and perhaps even should be discussed in the language classroom (Case, 2015; Clifford et al., 2013; Jolley & Maimone, 2015). The rationale for this potential integration varies, from a technological fatalism—MT is inevitable and therefore should be included in language pedagogy (Case, 2015)—to the need for pedagogical instruction to make student use of MT beneficial for learning (Clifford et al., 2013; Jolley & Maimone, 2015).

Research into perceptions about MT tools has also extended to students. Student users tend to view MT tools as helpful (Garcia & Pena, 2011; Lee, 2020; Niño, 2009; Tsai, 2019), although they also voice concerns about the accuracy of machine translation (Jin & Deifell, 2013; Jolley & Maimone, 2015; O'Neill, 2019; White & Heidrich, 2013). For instance, Jolley and Maimone (2015) noted a difference in perceived accuracy and text length, with students perceiving MT as better at translating shorter texts.

Research has also used survey instruments to gather information on how and why students use MT. Overall, students report frequent use of MT tools (Bourdais & Guichon, 2020; Clifford et al., 2013; Jin & Deifell, 2013; Jolley & Maimone, 2015; Larson-Guenette, 2013; O'Neill, 2019; White & Heidrich, 2013). Within this usage, students report relying on MT for a range of purposes (Bahri & Mahadi, 2016; Bourdais & Guichon, 2020; Clifford et al., 2013; Larson-Guenette, 2013; O'Neill, 2019). Commonly-reported uses were to look up words (e.g., vocabulary) (Clifford et al., 2013; Larson-Guenette, 2013; O'Neill, 2019; White & Heidrich, 2013) and to double check their work or instincts (Clifford et al., 2013; Jolley & Maimone, 2015). Another oft-cited motivation for choosing MT over other online resources relates to time: students report that MT's speed and efficiency is central in their decision to selecting these tools (Clifford et al., 2013; Larson-Guenette, 2013; O'Neill, 2019).

While there has been significant research on how students perceive and report using MT, much less has been done on how students actually use these tools. Calls to study how exactly students engage with online tools and platforms have been numerous (Chun, 2013; Fischer, 2007; Mroz, 2014). The importance of this gap is multifold. First, student perception data—how students perceive their own use—is not a fully accurate measure of how students actually use tools: past studies have shown discrepancies between

students' self-reported tool usage and actual usage (Fischer, 2007).

Second, how students use and engage with MT tools is a component of the larger ecology surrounding the intersection of MT and language teaching/learning. From an ecological theoretical perspective, multiple factors (e.g., experience, beliefs, platform design and functionality, policy) interact across scale levels (e.g., individual, classroom, institution, society) to impact how digital technologies are understood and used in language learning contexts (Blin, 2016; Kern, 2015). As demonstrated above, research has already documented one potential mediating factor—instructor beliefs—in how MT is understood and used in language learning contexts. However, another such factor—what students actually do with MT technologies—has been under-investigated in technology-enhanced language learning research and yet matters for our understanding of the impact of these tools on language teaching/learning.

Computer Tracking and L2 Writing

Computer tracking technologies—screen recording, eye trackers, keystroke and data logs—allow researchers to empirically observe how participants interact in technologically-mediated spaces and with technological tools, extending the researchers' purview beyond perception and belief data (Caws & Hamel, 2016; Hamel, 2012).

Past computer tracking studies have found widespread use of online resources in L2 writing tasks (Elola et al., 2008; Tight, 2017). Moreover, L2 writers have been shown to draw on online tools for complex and varied reasons (Elola et al., 2008; Lai & Chen, 2015; Tight, 2017; Yoon, 2016b, 2016a). For instance, Elola and colleagues (2008) found that undergraduate L2 learners of Spanish drew on online dictionaries for a range of reasons, from spelling to grammatical queries to refining writing style. More recently, Yoon (2016a) found that Korean graduate student L2 writers selected online tools for precise reasons based on perceived affordances and limitations of these tools.

Computer tracking studies have also shown that L2 writers' use of online tools is often idiosyncratic (Deifell, 2018; Yoon, 2016b) and that these diverse usage patterns make students more or less successful in their L2 writing (Caws et al., 2017; Elola et al., 2008; Yoon, 2016b). Elola and colleagues (2008) found that knowledge of dictionary abbreviations and looking for examples in online dictionaries led to better decisions. Yoon (2016b) showed that, over time, students reverted back to a reliance on familiar tools and patterns of interaction with said tools, despite previous training in more effective tools and uses.

 A few studies have examined how L2 writers use machine translation tools specifically. Garcia and Pena (2011) used screen capture technology and keystrokes to examine whether or not machine translation could support beginning learners' written composition. They found that students who were allowed to use MT paused less while writing than students who did not have access to MT. However, they also found that MT did not always lead to successful interventions or edits.

Both Tight (2017) and Deifell (2018) observed student use of a range of online tools, including machine translation. Tight (2017) found that undergraduate learners of Spanish (n = 12) made frequent use of platforms with online translation capabilities (e.g., Google Translate, SpanishDict), with relatively high success rates (88-87%). Using multiple computer-tracking technologies, (e.g., screen-capture, stimulated recall, and interviews), Deifell showed that L2 learners' (n = 2) use of online translators while writing was strategic, based on perceptions of the benefits and drawbacks of these tools, and intertwined with other tools and texts.

Current Article

The current article has two intertwined objectives. First, it looks to build on these past computer tracking studies on how students use MT tools, contributing additional research to the preliminary base described above. Second, this article looks to compare student use of MT tools with instructor perceptions. While survey-based studies have posed this question (e.g., Jolley & Maimone, 2015), a comparison between what students actually do with MT tools, garnered from computer-tracking-based methods, and instructor perceptions has not yet been undertaken.

Research Questions

1. How do second language learners use MT while writing?
2. How does student use of MT compare with instructor perceptions of MT?

Methods

This article draws on two sets of data: a survey study of educator beliefs about MT (Study 1) and a computer-tracking study on how university language learners of French and Spanish used MT (Study 2).

The two studies were related in several ways. First, both were created as a part of a larger inquiry into how MT impacts different scales of language education. Second, both studies look at university-level second language education in the US. That said, individual participants in the two studies were not related: Study 2 does not look at the MT practices of students taught by the instructors in Study 1.

Study 1: Second Educator Beliefs about MT

The first study used a qualitative survey to examine university second language educators' beliefs about machine translation in three areas: acceptability, student use and motivation, and the language teaching profession.

Context and Participants

The survey participants (*n* = 165) spanned three professional status categories: graduate students (25%), lecturers (30%), and professors (33%). A majority of participants (70%) taught languages commonly taught in the US (e.g., Spanish, French, German, Italian); representation from less commonly taught languages in the US (e.g., Chinese, Korean, Portuguese) was less strong (2% or less per language). Most participants (80%) taught at public four-year institutions, with less representation from smaller private institutions (17%). A total of 25 institutions participated from different US regions: Western US (*n* = 9); Southern US (*n* = 6); Eastern US (*n* = 5); Midwest US (*n* = 3); Southwest US (*n* = 2).

Data Collection

The survey used open-ended and closed-ended questions to investigate several areas of inquiry: acceptability, student use/motivation, and impact on the profession (Appendix A). The survey was developed by the co-authors in Fall 2018. To augment the validity of the instrument, definitions of MT were provided throughout the survey to increase the chances that key terms were understood similarly by participants (Lew et al., 2018). Reliability of the survey instrument was not assessed. The final version of the survey resulted from peer review and piloting with members of the target population (university FL instructors). The survey was distributed electronically in Spring 2019 via five national professional organizations (e.g., the Berkeley Language Center) and professional contacts.

Data Analysis

Closed-ended survey questions were analyzed using descriptive statistics. Open-ended survey questions were analyzed using open, inductive coding (Miles & Huberman, 1994; Saldaña, 2009) (Appendix B). Initial code categories were refined between the two researchers until application to 10% portions of the data reached at least 90% inter-rater reliability measures (Loewen & Plonsky, 2016). The full data set was analyzed by both authors using the finalized code list. All discrepancies in coding were noted, discussed, and resolved to increase rigor and reliability in the data analysis (Lew et al., 2018). All codes were then analyzed for their salience in the data and in relationship to each other in order to identify patterns across the data.

Study 2: Computer-Tracking Study on Student MT Use

The second set of data used in this article comes from a larger computer tracking study that examined how university language learners (French, Spanish, Mandarin) used online resources when writing. The current

article focuses on the French and Spanish language learners who participated in this larger project and their use of MT tools.

Context and Participants

Participants were recruited from second semester-level courses at three institutions in the United States: two large public universities and one community college. Two of these institutions (one public university, the community college) did not use placement exams; students elected into courses based on guidelines provided by the respective departments. At the third institution, the Spanish department used a placement exam while the French department offered (but did not require) a placement exam. An open invitation to participate in the research study was distributed via course websites and listservs. A total of 49 language learners completed study sessions (see Table 1 for demographic breakdown of participants). All study activities were completed via Zoom.

Table 1

Participant Summary (Study 2)

Language	French		Spanish		
	26		23		
Age	18-22	23-26	27-35	36-49	50+
	36	7	0	3	2
Gender	Female		Male		
	36		13		
Institution	Public University		Community College		
	40		9		
ACTFL Proficiency*	Novice		Intermediate		Advanced
	18		28		3

Note.* Participants' written proficiency level was assessed based on the writing task done in the study session (ACTFL Proficiency Guidelines, 2012).

During the research session, participants were asked to complete a short, written essay in the target language while their screen was recorded (Appendix C). This task was not connected to their language course, and it was made clear to students that it would not be graded or evaluated. Upon completion of the writing task, participants were asked to reflect on that task through a retrospective recall and post interview. The writing task was created to match students' proficiency level and to mirror activities that students would typically encounter at this level (Joyce, 1997, p. 59).

Data Collection

The primary data sources used in this study were: screen recordings of participants completing the writing task; a retrospective recall; and a post-interview.

Screen Recording. While participants completed the writing task, the session was recorded using Screencastify. Screen recording video totaled 680 minutes of task completion, with an average of 13:52 minutes per participant.

Retrospective Recall. A retrospective recall (also known as a stimulated recall or think aloud) asks participants to narrate a previously-completed task (Zhang & Zhang, 2020). This methodological technique enables researchers to gather information about language learner thought processes during a particular event (Gass & Mackey, 2016, p. 21) while mitigating the risks associated with simultaneous narration (e.g., cognitive overload) (Bowles, 2018; Zhang & Zhang, 2020).

To facilitate the narration of the task completion, several steps were taken. First, the retrospective recall was completed immediately after the task completion. Second, participants were trained to narrate that specific moment, as opposed to their current interpretation (Bowles, 2018; Gass & Mackey, 2016). Third, participants were shown specific moments from their task completion screen recording, to help them return to the moment of interest. (Full protocols are available on IRIS.)

These moments of interest were defined as "instances which expressed an immediate need on the part of the writer and his or her efforts to respond to a problem as identified through a series of visual signals in the screen recordings" (quoted from Hamel and Seror (2016, p. 145), in reference to a study by Park and Kinginger (2010)). For this study, these target moments hinged on leaving the composition document to use MT tools.

All retrospective recalls were video recorded via Zoom. The video files and their transcripts were used as the basis for analysis. These data totaled 764 minutes of video (with an average of 15:36 minutes per participant) and approximately 850 pages of transcripts.

Post-Interview. Semi-structured post-interviews were conducted following the retrospective recall and were used to triangulate the other data (Patton, 1990; Spradley, 1979). Broad categories included behavioral questions (what tools they use, how often, and how), attitudinal questions (tool preference, satisfaction with tools and results), and opportunities to follow-up on particular reactions or actions observed in the task completion or retrospective recall (Caws & Hamel, 2016; Hamel & Caws, 2010; Kuniavsky, 2012).

All post-interviews were video recorded via Zoom. Transcripts were used as the basis for analysis. The interviews produced 550 minutes of audio (with an average of 12 minutes per participant) and approximately 820 pages of transcripts.

Data Analysis

Student use data were coded using iterative rounds of coding (Miles & Huberman, 1994; Saldaña, 2009). Code categories focused on student actions (drawn from the task completion videos and triangulated with the retrospective recalls) and student thinking that accompanied actions (drawn from the retrospective recalls and post-interviews). For student actions, code application in the data analysis software (Dedoose) doubled as a way to quantify time spent engaged in different activities. The coding scheme used on the student use data can be found in Appendix D.

Like the attitudinal survey, code categories were refined between the two researchers and research assistants until application to 10% portions of the data reached at least 90% inter-rater reliability measures (Loewen & Plonsky, 2016). Discrepancies in coding were noted, discussed, and resolved to increase rigor and reliability in the data analysis (Lew et al., 2018). All codes were then analyzed for their salience in the data and in relationship to each other in order to identify patterns across the data.

Findings

Findings are organized around two axes of analysis: input—what is or should be put into MT tools—and output—what is done with the output of MT tools. The selection of input and output from the analytical battery stems from the article's overarching objective to compare student use and instructor perceptions: these analytical categories were chosen because we had appropriate data from each study to compare and because the juxtaposition of student use/instructor perceptions in these areas revealed important similarities or disconnects that will be fruitful for future pedagogical and research work.

Given the iterative and ecological nature of online tool use, in which input and output are often intertwined, the following findings sections are broadly organized around these categories, with some overlap between them.
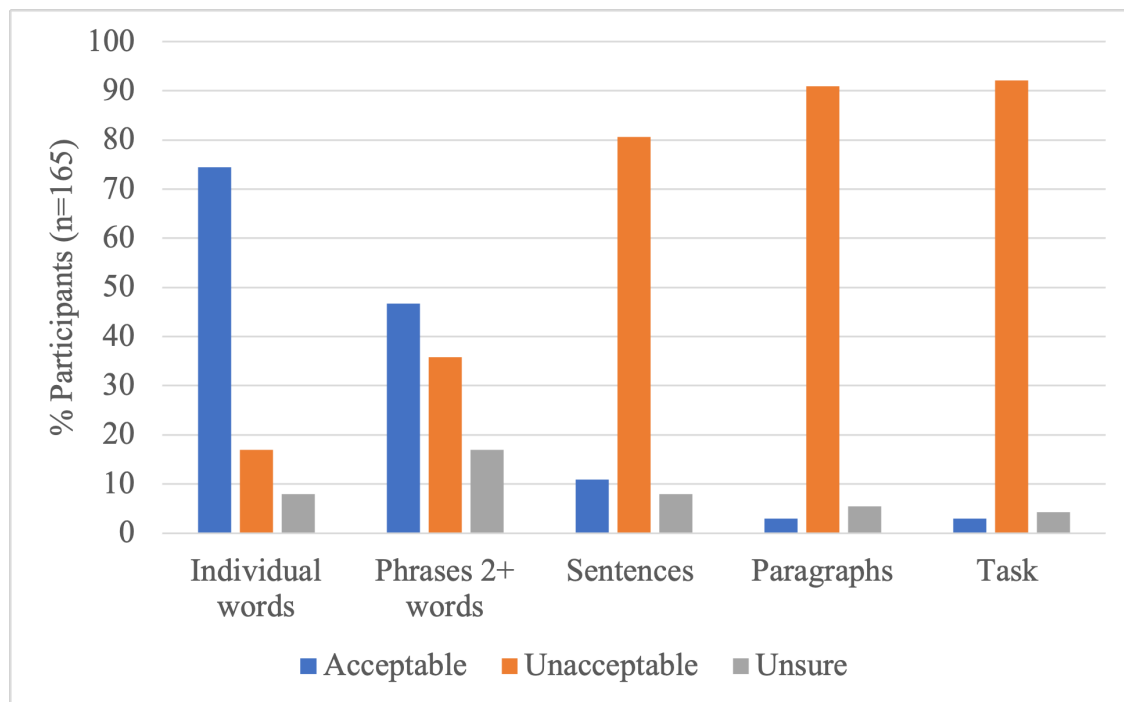
MT Input

Instructor Beliefs

Language educators surveyed had layered perceptions of both what students did with MT at the input level as well as what was acceptable. There was high consensus among educators that students use MT to look up words (95%), phrases (89%), and entire texts—specifically texts they had written in English, to be translated into the target language (82%). Fewer educators thought that students input entire texts they found on the internet (59%) or that they input texts they had written in the target language into MT, as a way to check their work (40%).

In parallel to these perceptions of what students put into MT tools, educators surveyed also drew clear lines around what kind of input they perceived as acceptable: MT was largely seen as appropriate for individual words and, to a lesser extent, phrases of 2+ words when writing (Figure 1). The acceptability dropped off significantly between the phrase and sentence level, which were across the board seen as unacceptable by teachers.

Figure 1

*Language Educator Perceptions of Acceptability of MT and Text Length for Writing*

For many instructors surveyed, these perceptions around acceptable MT input were reflected in their classroom policies around MT: of educators who reported having a policy on the use of MT (72%), 37% instituted partial bans of MT that prohibited the use of MT beyond individual words.

Student Use

There was a correspondence between instructor beliefs about what constitutes acceptable MT input and student use: as they completed the writing task, student participants most often used machine translation tools at the word level—that is, looking up individual words via MT tools (Figure 2). Input at the phrase level was the next most common type of input query among the population, followed by the sentence level (see Figure 3 for examples of each type of input, taken from student task completion video recordings).

Figure 2

Percentage of Input Queries
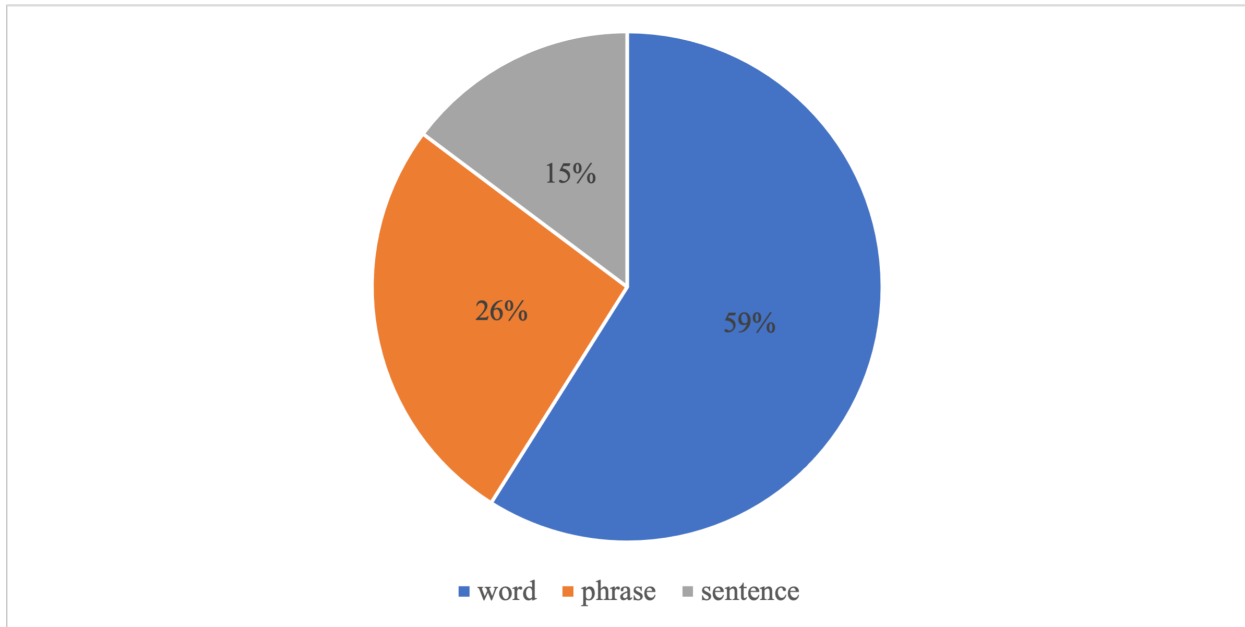
Looking more closely at the sentence level, few students input entire sentences in English and then fully copied the machine-generated translations into their composition documents (*n* = 8). There were cases of students who input entire sentences to clarify specific grammatical or syntactical questions, without copy/pasting the generated output into their composition documents. Others, as is illustrated in Figure 3.c, input full sentences they had written in the target language and checked their English translation. While a common strategy, this approach was not particularly successful in terms of catching errors or successfully negotiating meaning in student-composed texts.

Figure 3

Examples of Types of Input Queries (Word-level Input, Phrase-level Input, Sentence-level Input in the Target Language, and Sentence-level Input in English)

Figure 3.a

Figure 3.b

Figure 3.c

Figure 3.d
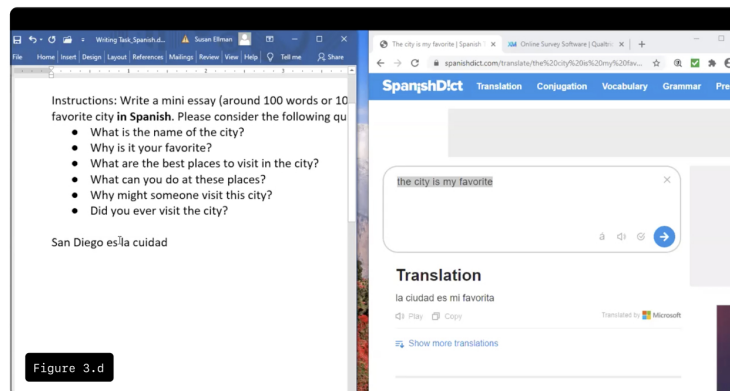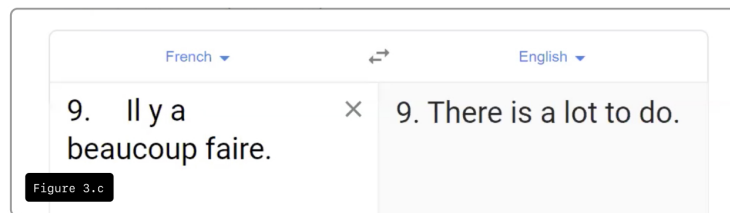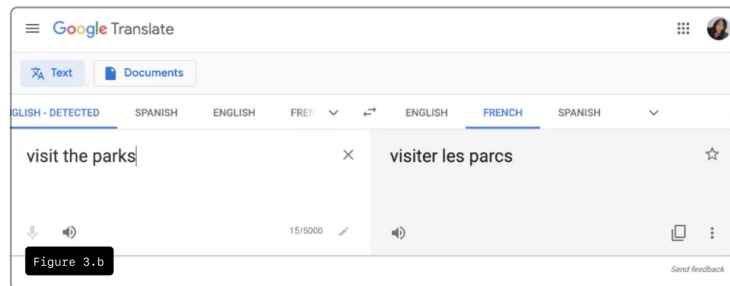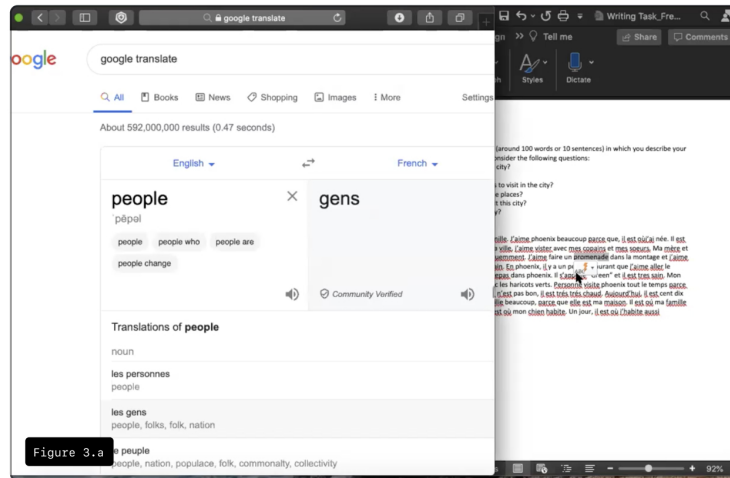
Importantly, our data suggest that the correspondence between teacher perceptions of acceptable MT use and how most students in our sample used MT was not always a coincidence. Rather, students were often influenced by their instructors' policies and beliefs about what was acceptable use of MT tools and fashioned their use of MT around these policies and beliefs. Indeed, approximately a third of students directly linked their use of MT to teacher policies. For instance, in her post-interview, Eliza explained her use of Google Translate this way:

> My teachers are always like "do not use [Google Translate]. It's not going to be right." So instead of even risking it, I always just try and do it like what I have learned. Because at the end of the day, [teachers] say, "you'd probably be—Like your grade would be punished more if you use Google Translate versus what I taught you, even if it's not correct." So that's, that's why I do it [use Google Translate] word for word.

Here, Eliza made a direct connection between how she used Google Translate—"word for word" or inputting individual words into MT tools—and how her instructors had framed MT tools and, more specifically, instructor beliefs around the acceptability of MT input. Eli made a similar connection, noting that his decision on whether and how to use Google Translate in his written task hinged on how his teacher had framed the tools: "In school, my teacher was like 'no you can use Google Translate but as a dictionary.'"

Using MT in this way—namely, for individual words—often had consequences for student writing, including incomplete or incorrect word selection and lack of agreement. For instance, without putting in more context, Meg's search for "place" in Google Translate led her to using a less appropriate word in her composition (Figure 4.a); alternatively, if Meg had put in more context, as in Figure 4.b, she would have been led to a more appropriate response.

Figure 4

*Illustration of Issues with Word-level Input*
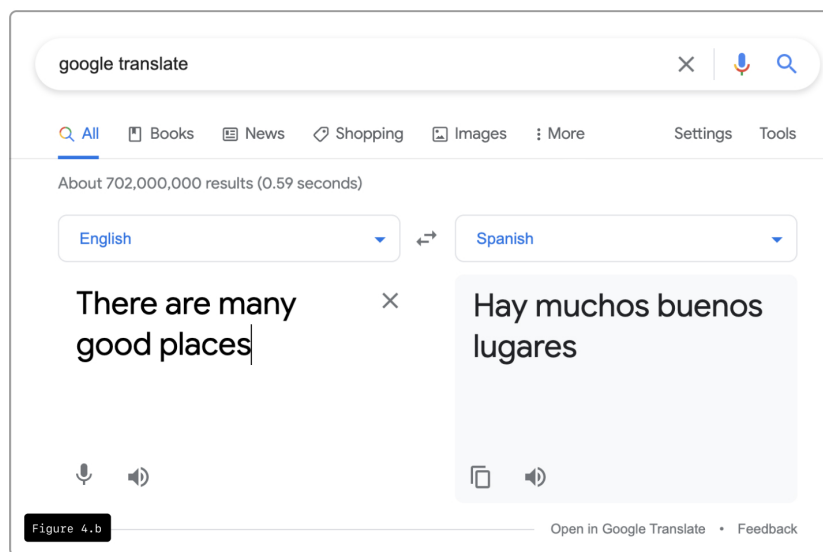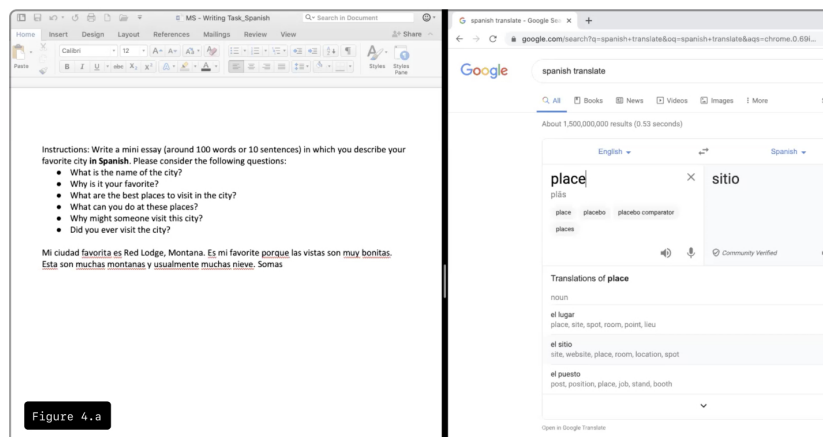
The impact of instructor policies on how students used MT was not limited to input but also extended to broader interactions with MT. Some students, for instance, fashioned their MT use to counteract any suspicions of "cheating," ensuring that they could justify what MT tools produced as their own knowledge. Sonia, for instance, was looking to describe the food in her favorite city. She initially typed "the food is delicious" into Google Translate (Figure 5a) but changed her original input to "the food is very good" (Figure 5b). In her retrospective recall, Sonia explained:

> I didn't know how to say delicious. And then I realized like, I don't know, my entire mindset whenever I'm using Google Translate is like, I can't use anything that like I should not have already learned, you know. I don't want to get into like trouble. So I make it really simple.

Importantly, this "trouble" was in reference to her instructors and the fear of accusations of violating the established rules around MT use.

Figure 5

Screenshots from Sonia's Task Completion

Figure 5.a

Figure 5.b
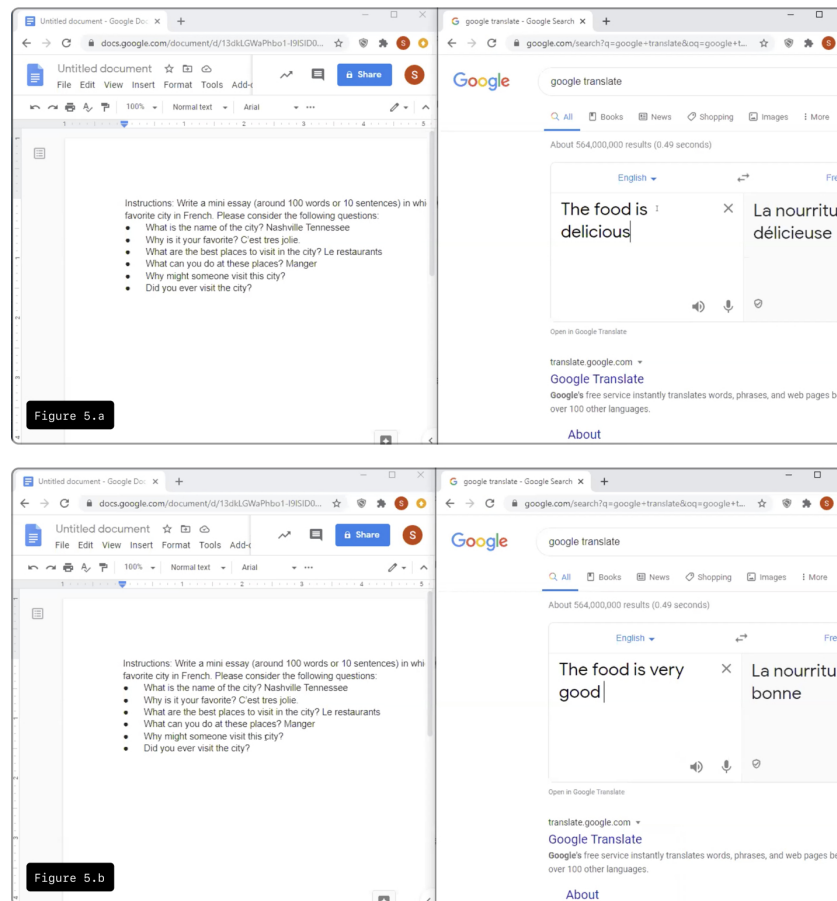
A related example comes from Kendrick. As he was typing "another fun thing to do" into Google Translate, he paused and replaced his search with "hiking." He explained that: "I thought it was cheating because it was like the whole sentence. And I didn't want to copy an entire sentence. So I just reverted back to what I could remember on the top of my head."

Other students linked their usage of Google Translate to grading policies. Gabby, for instance, described her general use of Google Translate as sticking to words she knows because "it is a more accurate representation of my knowledge at that time." However, she also admitted that she would have used Google Translate to look up that particular word if the assignment had been graded: "those fun vocabulary words make writing more interesting and more advanced and obviously that's better when you are trying to go for a better grade."

MT Output

Instructor Beliefs

The survey did not directly ask instructors what they believed students did with MT output. However, other questions included on the survey give some indication of what instructors believed. For instance, instructors were asked why they thought students used MT tools through closed-ended and open-ended questions. The most common response, cited by 70% of instructors surveyed, was that students use MT because it is quick, convenient, and fast, and because they are busy or unmotivated. One instructor put it this way:

> Sometimes you forget vocabulary words here and there (as a non-native speaker of Spanish, it happens to me too), but the point of a language class is for you to try to form sentences on your own. You can't learn and improve unless you make your own mistakes. If you just plug it into a translator, it's doing all the work for you, and you aren't taking the time to think critically about how you think you should form that sentence. You don't have to invest your energy into thinking about the mechanics of the language, you just spit something out that the translator gave you.

A prevalent belief amongst the instructors surveyed, in other words, was that MT reduced—or cut out altogether—the work, thinking, and energy required for language learning. This perception of MT being quick and easy suggests that many instructors surveyed believed that students do not get much from MT tool use, likely in part because instructors believe that they do not engage much with the output of MT tools.

Student Use

The computer-tracking study of student use of MT revealed that students' behavior with MT output both aligned with and diverged from instructor beliefs. One way to account for student actions with MT output is to quantify the instances in which students engaged or did not engage with the output of MT tools. This kind of engagement with MT output was operationalized in the coding in several ways, including presence or absence of copy/pasting MT output, pausing, scrolling through response possibilities, adjusting input, and cross-referencing.

Across the 49 participants, there were 302 instances in which students used MT tools. In about half these instances (n = 144), students did not review the output of MT; rather, they immediately copy/pasted or transferred the output to their composition documents. Alternatively, in the other half of these instances (n = 158), students reviewed the output of MT tools in some way. The average length of time that these students spent reviewing the MT output was 12 seconds, with a range of 2 to 71 seconds. Breaking this analysis down by student, 14 students did not or infrequently reviewed MT output; 10 students reviewed MT output a fourth to half the time; and 17 students reviewed MT output more than half of the time (see Table 2).

Table 2

Instances of MT Output Review by Student

Language	0-24%	25-49%	50-74%	75%-100%	No MT Use
French	6	6	6	5	3
Spanish	8	4	2	4	5
Total Students	14	10	8	9	8

Another way to account for student actions with MT output is to describe what they did with the output. We observed two primary categories of actions that students undertook to analyze MT output. One such action was seeking out and reviewing the different options, definitions, and examples of the target word or phrase provided by MT tools. This approach was seen in approximately one third of instances of output

analysis. Mailey, for instance, was looking for "a better word for 'area.'" When presented with Google Translate's output, she clicked to expand the search results offered and scrolled through them (Figure 6). In her retrospective recall, she described this effort to identify additional options and definitions in the following way: "Having all the options gives you the idea of how you want to say what you're trying to say next." In a similar vein, Allie described this approach this way: "I like to do that [look at the examples that online tools provide.]… I was just trying to read which one was what I was trying to say."

Figure 6

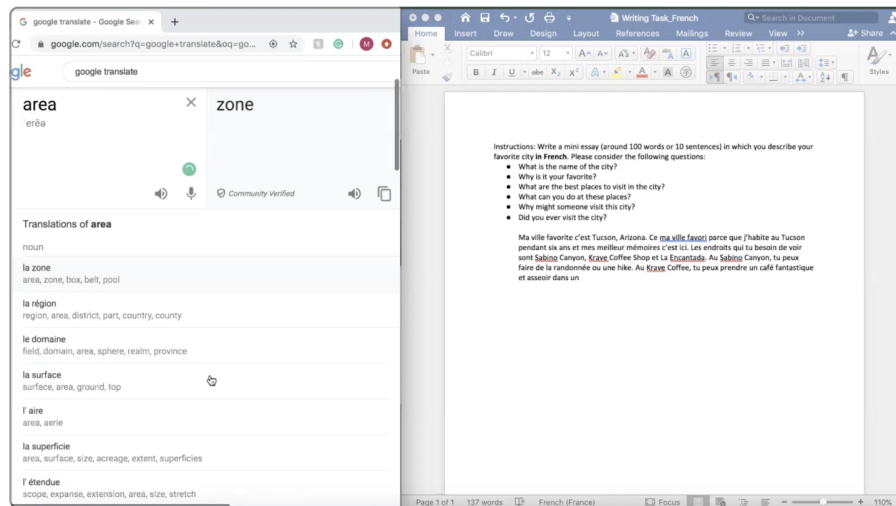*Screenshots from Mailey's Task Completion*

Figure 6

Another way that students in our sample engaged with MT output was re-checking—that is, expressly double checking or cross-referencing the output provided by MT tools. This approach was seen in approximately one fourth of the instances of output analysis. This took different forms, including looking up the English translation of the MT output result. A common strategy in this category was to simply switch the language of the input/output to see if the English translation for the original MT output made sense (single switch) or to see if the translation changed (double switch). Tim described this action this way: "a lot of times I will swap it between the two [languages] to see how Google changes it and then how it gets changed back but it's more like just confirming that it didn't do something really stupid."

Another common way that student participants re-checked the output of MT was cross-referencing the search output with another query, either in the same MT tool or in another tool (MT or otherwise). Carlana, for instance, began her search for "dish" (a regional type of food) with Google Translate (Figure 7.a). She then altered the input a few times (Figure 7.b), explaining in her retrospective recall that

> I would like to find something in context. This [the initial output provided by Google Translate] could work but I'm not sure it would work and I won't really know if it would work unless I see it in a sort of a context I guess. So I was like "ok I'm looking"—I think I was looking to see if there were any different terms that would pop up. More explanation of what they meant.

Carlana then switched to Cambridge French-English dictionary because, she reflected after the task, the dictionary "would probably have more definitions of the word I was looking for." She scrolled through the entries and reviewed the dictionary's definition and examples of the initial output provided by Google Translate (Figure 7.c). This process made Carlana more comfortable selecting the word provided by Google Translate: "the definitions were clear and there was more context to go off of."

Figure 7

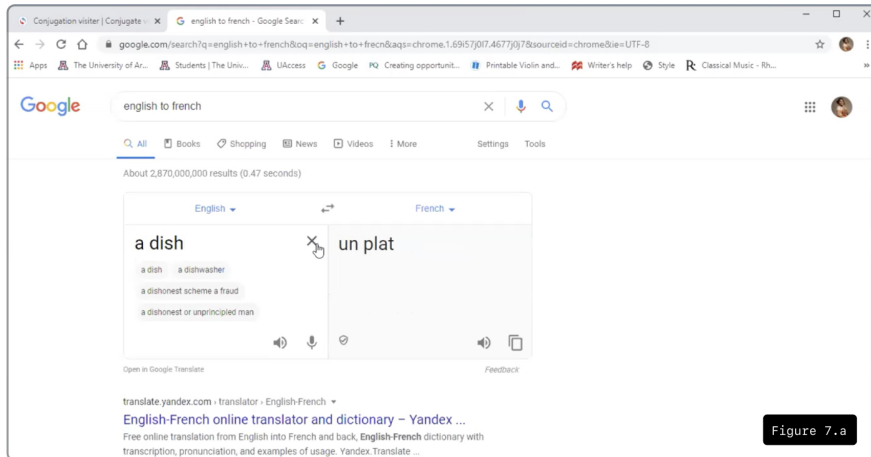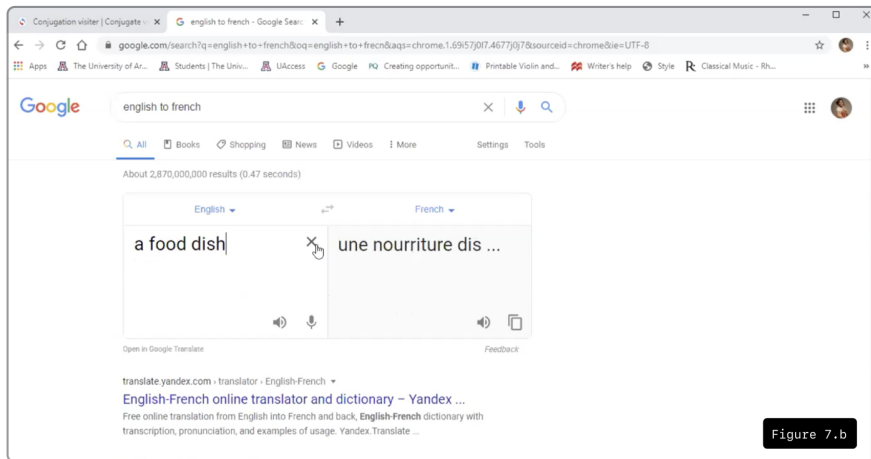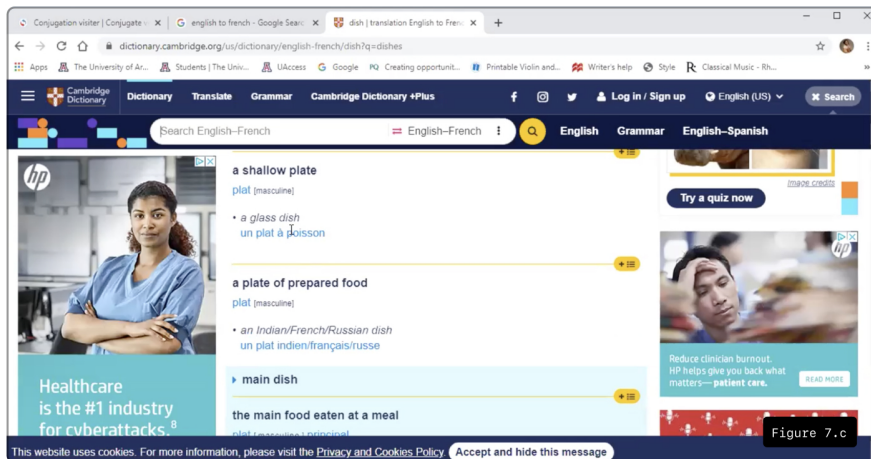*Screenshots from Carlana's Task Completion*

Discussion

This article asked how second language learners use MT for writing in terms of input/output and how this use compared with instructor perceptions. In terms of input, students most often used MT to look up single words, usage that was expected and approved of by instructors. Conversely, students less often input whole sentences or texts they had written in English into MT tools, as was predicted by instructors. Importantly, student use of MT was often influenced by instructor policies and perceptions.

In terms of output, while some student participants took MT output at face value—that is, they used MT results without question—, many often engaged with the output of MT tools using complex strategies. This usage contrasted with instructor expectations of largely uncritical student use of MT.

These findings confirm some reports from past survey-based studies of student MT use, including the high frequency of MT tool use (Bourdais & Guichon, 2020; Clifford et al., 2013; Jin & Deifell, 2013; Jolley & Maimone, 2015; Larson-Guenette, 2013; O'Neill, 2019; White & Heidrich, 2013) and how learners often use MT—to look up words (Clifford et al., 2013; Larson-Guenette, 2013; O'Neill, 2019; White & Heidrich, 2013) and to double check their work (Clifford et al., 2013; Jolley & Maimone, 2015).

The look at precisely how students used MT provided by the computer tracking study provided additional insight into these actions. For instance, several ways in which students double checked their work were identified, including surveying options and examples provided by MT, cross-referencing MT output, and back translating their target-language texts into English. These strategies dovetail with the diverse and complex patterns of use observed in other computer-tracking studies of other online tools (Caws et al., 2017; Elola et al., 2008; Yoon, 2016b).

Juxtaposing this student use with instructor perceptions highlights that student use of MT can be, but is not always, more complex than instructors realize. The thoughtfulness and thoroughness that was observed in some student practices has implications for how to approach MT in the classroom, as discussed in the next section.

In addition, the findings of this study highlight that instructor policies, undergirded by beliefs about MT, can mediate what students do with MT and have ramifications for how successful they can be. For instance, instructor preference for word-level inputs into MT encourages students to use MT tools in ways that are not in-line with current functionalities of MT tools: current versions of machine translation require lengthier inputs to more accurately identify the "context" of the query, defined in machine learning as proximity to the target word in the large data bases of language that undergird MT platforms (Poibeau, 2017). In other words, instructor policies and student practices collide with an additional mediating factor, MT tools themselves.

Another potential ramification of the mediation of instructor MT policies and student MT use connects to learning, specifically in relation to student fear over cheating. While student reticence to use unfamiliar MT output might signal a successful MT policy to some, we would argue that this approach can limit learning. For instance, in the previous example from Sonia, the word "delicious" is the cognate in French (délicieux, -euse). Rather than using and potentially learning this new word, however, Sonia erred on the side of caution and replaced it with a simpler input that wouldn't put her in danger of getting "in trouble." In our view, overemphasizing cheating and MT could, in other words, restrict student MT use in ways that also restricts their learning.

Pedagogical Implications

These findings suggest several implications for practice. First, the findings point to the importance of interrogating our own beliefs about what students do with MT tools. Our students can surprise us and teach us what they need in terms of MT use and language learning more broadly. In a classroom context, identifying how students use MT—and, by extension, what students need in terms of MT use and support— could come from asking students as well as informal observations of how they use MT in classroom-based

activities.

Second, the layered mediating factors that are showcased in this article point to the need to move beyond teaching about MT in limited ways, such as banning it, pointing out its inaccuracies, or restricting its use to words. While these approaches to MT training and policies have been the norm, they are not supporting students in meaningful ways and, in some cases, are actually hindering them. Rather, training and policies around MT and other online tools should look to respond to students' needs and further support the work that they are already doing by teaching them to use MT more critically and for language learning.

For example, it would be beneficial to bolster students' strategies for analyzing the output of MT and other online tools. For instance, instructors could select words/phrases in English that are (a) beyond students' level, and (b) have multiple definitions in the target language. Possible examples for the learners in this article could be: "hike" or "might" (Spanish) and "model" or "fly" (French). With these entries, instructors could model how to seek out the context or additional information needed to make informed decisions. Students could then be given additional words/phrases to research independently, so as to practice these strategies on their own.

It would also be useful to teach students what types of input into MT tools are most likely to lead them to more appropriate answers. This could be done by showing students how very similar searches can produce very different results within MT tools. For learners of French, for instance, instructors could showcase the different results of "I fly" (*je vole*), "fly" (*voler*), "a fly" (*une mouche*), and "flying" (*en volant*).

This last point relates to a larger component of training on MT highlighted by the findings of this study: the need to augment student awareness of how MT tools work given that the innerworkings of MT tools mediate student use. For instance, when teaching students how to manipulate MT input to their advantage, instructors could reference how MT tools, powered by algorithms, produce results based on context and probability—the likelihood of a particular meaning or translation is based on the words included in the original entry. For example, MT tools are more likely to correctly translate "I fly" than "fly" because the former contains the context to disambiguate the verb from the insect. Even without an extensive foray into the mechanics of neural networks or machine translation, this kind of information should clarify for students the how and the why of manipulating MT input for more accurate translations.

Broadly teaching students how MT tools work in this way, combined with training for language learning specifically, would likely make students more savvy users of these tools and further increase their ability to successfully manipulate them for language learning. This approach dovetails with other pedagogical approaches in technology-mediated language learning that advocate for the denaturalization of technology: the seamless prevalence of technological tools in our daily lives has led to their normalization, which engenders a lack of attention to the mediating factor technologies play in various personal and professional domains (Darvin, 2017; Hellmich, 2019; Jones, 2019; Kern, 2015).

Limitations and Future Research

Future studies should continue to prioritize work that examines precisely *how* students use MT tools to further push our understanding and preconceived beliefs about how students use these tools. While more time-intensive than survey-based studies, observations of student use are essential in developing ways to effectively address tools like MT in language teaching and learning. Future studies might take advantage of well-established computer-tracking methodologies in CALL (e.g., eye tracking, Smith, 2009) or recent advances in CALL methodologies (e.g., Smith, 2017) to add additional nuance to how students engage with online tools, input and output. In addition, studies that look at the impact of MT use on language learning are necessary (Jolley & Maimone, in press).

There are limitations of this study that could be used to launch future studies. First, the analysis of student use of machine translation was not differentiated based on proficiency level. It would be beneficial in future research studies to explore how student use of MT tools varies based on proficiency level as well as other factors (e.g., individual languages, less-commonly taught languages, heritage vs. non-heritage language

learners). In addition, the study focused on a singular task completed in an artificial environment. Future work that focuses on capturing student use of MT tools in more naturalistic settings—perhaps screen recordings of writing tasks completed at home—and over time would help to expand understandings of student use of MT. Finally, the ideas for new pedagogical interventions discussed above would be an important additional area of future research: examining how different training approaches impact student use of MT tools.

Conclusion

Juxtaposing student use and instructor perceptions has revealed that, rather than leveraging MT tools uncritically, students can display complex ways of engaging with MT that signal attempts to use it thoughtfully within the parameters set out by instructors and the MT tools themselves. However, the complex task of negotiating these instructor parameters and the MT tools themselves also leads to problems. This study points to the necessity of continued dialogue between instructors and students through further research and the development of classroom-based activities. Ultimately, through such dialogue, students and instructors can co-construct classroom policies that invite considerations of the strategies students are already using and their effectiveness, provide training in the functionality and limits of the tools vis-à-vis these strategies, and, ultimately, support MT use in ways that potentially open new opportunities for learning.

Acknowledgements

References

ACTFL Proficiency Guidelines. (2012). American Council for the Teaching of Foreign Languages. http://www.actfl.org/publications/guidelines-and-manuals/actfl-proficiency-guidelines-2012

Barr, D. (2013). Embedding technology in translation teaching: Evaluative considerations for courseware integration. *Computer Assisted Language Learning*, *26*(4), 295–310. https://doi.org/10.1080/09588221.2012.658406

Blin, F. (2016). Towards an 'ecological' CALL theory: Theoretical perspectives and their instantiation in CALL research and practice. In F. Farr & L. Murray (Eds.), *The Routledge handbook of language learning and technology* (pp. 39–54). Routledge. https://doi.org/10.4324/9781315657899

Bourdais, A., & Guichon, N. (2020). Représentations et usages du traducteur en ligne par les lycéens. *Alsic*, *23*(1). https://doi.org/10.4000/alsic.4533

Bowles, M. A. (2018). Introspective verbal reports: Think-alouds and stimulated recalls. In A. Phakiti, P. de Costa, L. Plonsky, & S. Starfield (Eds.), *The Palgrave handbook of applied linguistics research methodology* (pp. 339–357). Palgrave Macmillan UK. https://doi.org/10.1057/978-1-137-59900-1

Briggs, N. (2018). Neural machine translation tools in the language learning classroom: Students' use, perceptions, and analyses. *JALT CALL Journal*, *14*(1), 3–24.

Case, M. (2015). Machine translation and the disruption of foreign language learning activities. *ELearning Papers*, *45*, 4–16. http://urn.kb.se/resolve?urn=urn:nbn:se:du-20314

Caws, C., & Hamel, M. J. (2016). *Language-learner computer interactions: Theory, methodology and CALL applications*. John Benjamins.

Caws, C., Leger, C., & Perry, B. (2017). Peer editing in French using digital tools: A micro-analysis of learner-computer interactions. *Canadian Journal of Applied Linguistics*, *20*(1), 77–97. https://journals.lib.unb.ca/index.php/CJAL/article/view/24491

Chun, D. M. (2013). Contributions of tracking user behavior to SLA research. *CALICO Journal*, *30*, 256–262. https://www.jstor.org/stable/calicojournal.30.256

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

Correa, M. (2011). Academic dishonesty in the second language classroom: Instructors ' perspectives. *Modern Journal of Language Teaching Methods*, *1*(1), 65–80.

Darvin, R. (2017). Language, power, and critical digital literacy. In S. Thorne & S. May (Eds.), *Language, education and technology* (pp. 17–30). Springer.

Deifell, E. D. (2018). Dynamic intertextuality and emergent second language microdevelopment in digital space. [Doctoral dissertation, University of Iowa]. Proquest.

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

Elola, I., Rodriguez-Garcia, V., & Winfrey, K. (2008). Dictionary use and vocabulary choices in L2 writing. *Estudios de Linguistica Inglesa Aplicada*, *8*, 63–89.

Fischer, R. (2007). How do we know what students are actually doing? Monitoring students' behavior in CALL. *Computer Assisted Language Learning*, *20*(5), 409–442. https://doi.org/10.1080/09588220701746013

Garcia, I., & Pena, M. I. (2011). Machine translation-assisted language learning: Writing for beginners. *Computer Assisted Language Learning*, *24*(5), 471–487. https://doi.org/10.1080/09588221.2011.582687

Gass, S. M., & Mackey, A. (2016). *Stimulated recall methodology in applied linguistics and L2 research*. Routledge.

Hamel, M. J. (2012). Testing aspects of the usability of an online learner dictionary prototype: A product- and process-oriented study. *Computer Assisted Language Learning*, *25*(4), 339–365. https://doi.org/10.1080/09588221.2011.591805

Hamel, M. J., & Caws, C. (2010). Usability tests in call development: Pilot studies in the context of the dire autrement and francotoile projects. *CALICO Journal*, *27*(3), 491–504. https://doi.org/10.11139/cj.27.3.491-504

Hamel, M. J., & Seror, J. (2016). Video screen capture to document and scaffold the L2 writing process. In C. Caws & M. J. Hamel (Eds.), *Language-learner computer interactions: Theory, methodology and CALL applications* (pp. 137–162). John Benjamins.

Hellmich, E. A. (2019). A critical look at the bigger picture: Macro-level discourses of language & technology in the US. *CALICO Journal*, *36*(1), 39–58. https://doi.org/10.1558/cj.35022

Hellmich, E. A., & Vinall, K. (2021). FL instructor beliefs about machine translation: Ecological insights to guide research and practice. *International Journal of Computer-Assisted Language Learning & Teaching*, *11*(4), 1–18. https://doi.org/10.4018/IJCALLT.2021100101

Jin, L., & Deifell, E. (2013). Foreign language learners' use and perception of online dictionaries: A survey study. *MERLOT Journal of Online Learning and Teaching*, *9*(4), 515–533.

Jolley, J. & Maimone, L. (2015). Free online machine translation: Use and perceptions by Spanish students and instructors. In A. J. Moeller (Ed.), *Learn Languages, Explore Cultures, Transform Lives* (pp. 181–200). Central States Conference on the Teaching of Foreign Languages.

Jones, R. H. (2019). The text is reading you: Teaching language in the age of the algorithm. *Linguistics and Education*, *62*. https://doi.org/10.1016/j.linged.2019.100750

Joyce, E. (1997). Which words should be glossed in L2 reading materials? A study of first, second, and third semester French students' recall. *Pennsylvania Language Forum*, 58–64.

Kelleher, J. D. (2019). *Deep learning*. MIT Press.

Kern, R. (2015). *Language, technology, and literacy*. Cambridge University Press.

Kuniavsky, M. (2012). *Observing the user experience: A practitioner's guide to user research*. Morgan Kaufmann.

Lai, S. L., & Chen, H. J. H. (2015). Dictionaries vs concordancers: Actual practice of the two different tools in EFL writing. *Computer Assisted Language Learning*, *28*(4), 341–363. https://doi.org/10.1080/09588221.2013.839567

Larson-Guenette, J. (2013). "It's just reflex now": German language learners' use of online resources. *Die Unterrichtspraxis/Teaching German*, *46*(1), 62–74. https://doi.org/10.1111/tger.10129

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

Lew, S., Yang, A. H., & Harklau, L. (2018). Qualitative methodology. In A. Phakiti, P. De Costa, L. Plonsky, & S. Starfield (Eds.), *The Palgrave handbook of applied linguistics research methodology* (pp. 79–101). Palgrave Macmillan.

Lewis-Kraus, G. (2016, December 14). The great A.I. awakening. *The New York Times*. https://www.nytimes.com/2016/12/14/magazine/the-great-ai-awakening.html

Loewen, S., & Plonsky, L. (2016). *An A-Z of applied linguistics research methods*. Palgrave.

Miles, M. B., & Huberman, A. M. (1994). *Qualitative data analysis: An expanded sourcebook* (2nd ed.). Sage Publications.

Mroz, A. P. (2014). Process research screen capture. In C. A. Chapelle (Ed.), *The encyclopedia of applied linguistics* (pp. 1–7). John Wiley & Sons, Ltd. https://doi.org/10.1002/9781405198431.wbeal1450

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

O'Neill, E. M. (2019). Online translator, dictionary, and search engine use among L2 students. *CALL-EJ*, *20*(1), 154–177.

Park, K., & Kinginger, C. (2010). Writing/thinking in real time: Digital video and corpus query analysis. *Language Learning & Technology*, *14*(3), 31–50. https://hdl.handle.net/10125/44225

Patton, M. Q. (1990). *Qualitative evaluation and research methods*. Sage Publications.

Poibeau, T. (2017). *Machine translation*. MIT Press.

Qun, L., & Xiaojun, Z. (2015). Machine translation: General. In C. Sin-wai (Ed.), *The Routledge encyclopedia of translation technology* (pp. 105–119). Routledge. https://doi.org/10.5860/choice.188800

Saldaña, J. (2009). *The coding manual for qualitative researchers* (2nd ed.). Sage Publications.

Smith, B. (2009). The relationship between scrolling, negotiation, and self-initiated self-repair in an SCMC environment. *CALICO Journal*, *26*(2), 231–245. https://doi.org/10.1558/cj.v26i2.231-245

Smith, B. (2017). Commentary: Methodological innovation in CALL research. *Language Learning & Technology*, *21*(1), 1–3. https://dx.doi.org/10125/44591

Spradley, J. P. (1979). *The ethnographic interview*. Wadsworth.

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

Tight, D. G. (2017). Tool usage and effectiveness among L2 Spanish computer writers. *Estudios de Lingüística Inglesa Aplicada*, *17*, 157–182. http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=19893492

Tsai, S. (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

White, K. D., & Heidrich, E. (2013). Our policies, their text: German language students' strategies with and beliefs about web-based machine translation. *Die Unterrichtspraxis/Teaching German*, *46*(2), 230–250. https://doi.org/10.1111/tger.10143

Wu, Y., Schuster, M., Chen, Z., Le, Q. V., Norouzi, M., Macherey, W., Krikun, M., Cao, Y., Gao, Q., Macherey, K., Klingner, J., Shah, A., Johnson, M., Liu, X., Kaiser, Ł., Gouws, S., Kato, Y., Kudo, T., Kazawa, H., … Dean, J. (2016). *Google's neural machine translation system: Bridging the gap between human and machine translation*. https://arxiv.org/abs/1609.08144

Yoon, C. (2016a). Concordancers and dictionaries as problem-solving tools for ESL academic writing. *Language Learning & Technology*, *20*(1), 209–229. https://hdl.handle.net/10125/44453

Yoon, C. (2016b). Individual differences in online reference resource consultation: Case studies of Korean ESL graduate writers. *Journal of Second Language Writing*, *32*, 67–80. https://doi.org/10.1016/j.jslw.2016.04.002

Zhang, L. J., & Zhang, D. (2020). Think-aloud protocols. In J. McKinley & H. Rose (Eds.), *The Routledge handbook of research methods in applied linguistics* (pp. 302–311). Routledge.

Appendix A. Study 1 Survey Questions

Area of Inquiry	Open-Ended Questions	Closed-Ended Questions
Acceptability	Do you think that it is acceptable to use MT for some tasks and not others?	Students are completing a <u>writing*</u> task in the target language. Students are using MT to translate from what they have personally written into the target language. Assess the appropriateness of the following uses of MT for language students: individual words, phrases, complete sentences, entire paragraphs, entire tasks. *Acceptable, Unacceptable, Unsure* *Scenarios were asked for four skills: reading, writing, listening, speaking. How much do you agree or disagree with the following statement: "My assessment of the appropriateness of using MT for language learning depends on whether the task is graded or ungraded.*"* *Five-point Likert scale from Strongly agree to Strongly disagree*
	If so, what is this policy? Feel free to summarize or copy/paste from the policy document.	Do you have a policy on students using Machine Translation (MT) to complete tasks in your class? *Yes/No/N/A*
Student Use & Motivation	Why do you think students use MT? What, if anything, do students learn through the use of MT?	Below are common reasons for student use of MT. In your opinion, why do students use MT? Select the top four reasons · *They don't want to do the work themselves* · *They don't have time to complete the task* · *They don't understand the material and/or task* · *They are behind in their overall work for the class*

		· *They don't have the required proficiency-level to complete the task* · *They are worried about their grade* · *They are using all resources available to them* · *They want to learn faster than the course pace* · *They lack confidence in their own language abilities* · *They want to convey their ideas in the L2* · *Other (please specify)* From the following list, select all the ways that you imagine students <u>at your institution</u> are using MT. · *Looking up words* · *Translating isolated phrases from the target language to English/student's primary language* · *Translating their own texts, written in English/student's primary language, into target language* · *Translating text found on Internet into target language* · *Translating their own texts, written in target language, back into English/student's primary language* · *Other (please specify)* Estimate the proportion of students whom you believe use Machine Translation (MT) to complete any course tasks. *None, 10%, 25%, 50%, 75%, 100%*
Profession	Please explain your answer to the last question (How much do you agree or disagree with the following statement: "MT threatens the language teaching profession.").	How much do you agree or disagree with the following statement: "MT threatens the language teaching profession." *Five-point Likert scale from Strongly agree to Strongly disagree*

Appendix B. Study 1 Codebook

ACCEPTABILITY, LEARN	
Code	*Description*
Cheat	MT is not acceptable, constitutes academic dishonesty
Depends on	Acceptability or learning depends on different factors, such as
-context	Where use of MT happens
-course	Type of course
-level	Language learner level
-skill	4 skills
-task	Type or difficulty of task
-student	Type of student
Limitations of MT	Students learn about the limitations of MT
Language	Students learn linguistic elements of language (e.g., syntax, grammar, vocab)
Nothing	Students learn nothing from MT
STUDENT USE & MOTIVATION	
Code	*Description*
Cheat	Students use MT expressly for academic dishonesty.
Busy	Students use MT because they are busy, pressed for time.
Lazy	Students use MT because they are lazy/not willing to put in the time/effort to learn.
Quick	Students use MT because it's easily accessible, convenient, quick. Includes saving time.
Learning	Students use MT to help support language learning.
Grades	Students use MT due to concern over grades.
Programmatic	Students use MT to satisfy programmatic reasons, such as language requirement
Confidence	Students use MT due to a lack of confidence in language abilities.
Preparation	Students use MT because they are not prepared for course requirements.

PROFESSION	
Code	*Description*
Communication	MT is not a threat because it does not teach communication.
Human	MT is not human and therefore cannot threaten the profession.
Instructor	MT does not replace what instructors do
Skills	MT does not support skill building/learning
Language in the US	Broader ideologies of language in the US threaten profession
Language use	The ways students want to use language (fulfill requirement, travel) is a threat to the profession.

Appendix C. Writing Task (Study 2)

Instructions: Write a mini essay (around 100 words or 10 sentences) in which you describe your favorite city **[in French/in Spanish]**. Please consider the following questions:

- What is the name of the city?
- Why is it your favorite?
- What are the best places to visit in the city?
- What can you do at these places?
- Why might someone visit this city?
- Did you ever visit the city?

Appendix D. Study 2 Codebook

TOOL USE			
ACTIONS Tool(s) student uses to complete the task		**THINKING** The cognitive rationale behind tool use	
Code	*Description*	*Code*	*Description*
Multiple Tools	Student uses multiple tools or sites in completing one transaction	Intentional use	Student expresses using a tool for a specific reason that aligns with their intention/what they are looking for
Notes/ Textbook	Student uses notes/textbook	Confidence	Students express lack of confidence in their language as motivating tool use. Explicit reference to doubt.
Online Tools	Student uses online tools (subcodes: GoogleTranslate, Reverso, WordReference, SpanishDict, Other)	Efficiency	Students explains tool use (or non-use) in terms of time, accessibility, efficiency
		Not Satisfied	Student reports tool use because they weren't happy with what they had written
		Intentional non-use	Student reports intentionally not going to a tool
TOOL INPUT			
ACTIONS What students put into the tool		**THINKING** The cognitive rationale behind tool input	
Code	*Description*	*Code*	*Description*
Look Up	Student reports looking something up that they didn't know (subcodes: word, phrase, sentence, spelling, gender, verb form, information)	Intentional input	Student reports intentionally putting in more or less into tool so as to get at a better translation
Verify	Student reports thinking they know but not being sure (subcodes: word, phrase, sentence, spelling, gender, verb form, information)		
Text input	Length of text student inputs into MT tools based on task completion (subcodes: word, phrase, sentence)		
Compose	Student writes more than 1 sentence in the L1 or L2		
Change input	Student changes input		

TOOL OUTPUT			
ACTIONS What students do with the output of the online tool		**THINKING** The cognitive rationale behind tool output manipulation	
Code	*Description*	*Code*	*Description*
Tool Output	Time spent on tool output	Familiar	Student accepts/rejects output because they mention something either looking familiar or not looking familiar.
Analyze Output	Student reports analyzing the output of the tool. Subcodes: examples (seeking out examples), re-checking (confirming output)	Satisfaction	Student accepts/rejects output because they are satisfied or dissatisfied (like/don't like). Broader than familiar.
Rephrase	Student changes or rewrites output	Match Output	Student accepts or rejects output (or edits output) because it does or doesn't match their language level
No analysis	No apparent analysis or rationale for use of output		
Change course	Without or without analysis, student states that they reject output. This may include rephrasing their initial idea in the text document or abandoning their idea altogether.		

About the Authors

Emily A. Hellmich (PhD, University of California, Berkeley) is Associate Director of the Berkeley Language Center (University of California, Berkeley). Her work focuses on the impact of our global, digital world on language education. Her research has appeared in *Teachers College Record*, *CALICO*, *Foreign Language Annals*, and *Computer Assisted Language Learning*.

E-mail: eahellmich@berkeley.edu

Kimberly Vinall (PhD, University of California, Berkeley) is the Executive Director of the Berkeley Language Center (University of California, Berkeley). She explores how learners' and teachers' understandings of cultural/linguistic differences are mediated by identity constructs, teaching materials, digital tools, and ideologies. She has published in the fields of applied linguistics, CALL, and SLA.

E-mail: kvinall@berkeley.edu