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

JASON R. JOLLEY

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

LUCIANE MAIMONE

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

Although use of machine translation (MT) technologies by learners may seem like a relatively new issue in foreign language (FL) education, researchers have been investigating connections between MT tools and FL teaching and learning for more than three decades, years before learners had access to free online services such as Google Translate. This literature review summarizes this rapidly expanding research domain both chronologically and thematically, identifying key concepts, insights, and findings and mapping them onto a framework inspired by questions commonly asked by both researchers and practitioners: How do learners use MT tools? What do instructors and learners think about MT? How might MT use affect language learning? How should instructors respond to MT use by learners? By summarizing and drawing connections between the assumptions, methods, and findings of key studies in these categories, this review provides a historical perspective and suggests new directions for future research.

INTRODUCTION

The collision course between machine translation (MT) and the field of language education has been accelerating for the past 10 to 15 years, as ever smarter and smaller devices, coupled with ubiquitous cellular and Wi-Fi connectivity, have put free online MT websites and apps at learners' fingertips. Understandably, many instructors are skeptical about the quality of MT output and are quick to equate student use of MT with academic dishonesty. Others, aware of the excesses of the Grammar-Translation Method (GTM), eschew translation tasks altogether, even as evidence mounts for a reevaluation of first language (L1) use, and translation specifically, in the second language (L2)[1] classroom (see, e.g., Atkinson, 1993; G. Cook, 2010; V. Cook, 2001; Kerr, 2014; Osswald, 2010; Ramsden, 2018; Vermes, 2010). Given the recent nature of these developments, the fact that many practitioners—and even researchers—assume that little is known about the intersection of MT and language teaching and learning is understandable. However, researchers have been exploring the implications of MT for language teaching and learning for more than 30 years. After briefly discussing advances in MT technology and accessibility and identifying five major strands in the broader MT literature, we outline key concepts, findings, and implications from MT studies relevant to language teaching and learning published in the past four decades.

Although a comprehensive history of the evolution of MT technologies is beyond the scope of this review,[2] readers should be familiar with major developments in the field, since advances in MT system capabilities and accessibility affect a number of concerns addressed in the literature, including learner and instructor beliefs, the ability to detect unauthorized MT use, and strategies for reacting to it. Hutchins (2010) traces interest in the development of modern MT systems to the late 1940s, coinciding with the invention of computers. Research into automatic translation continued during the technology boom of the Cold War years, driven in part by major advances in linguistics. However, it was not until the 1970s and 1980s that MT systems such as SYSTRAN, Logos, and METAL were used commercially, though their accessibility was hampered by the fact that they had to be run on large mainframes, often in government or military installations. As computing went mainstream in the 1980s and 1990s, some of these products were adapted for use on desktop workstations, making them more readily accessible in academic settings, including in language labs.

The technology driving most of these early systems was rule-based MT, which uses dictionary-based lexical substitutions and pre-programmed morphosyntactic rules to perform automated transfer operations. In subsequent decades, this transfer approach would be supplanted by increasingly sophisticated, better-performing, and more accessible technologies. These include example- and phrase-based statistical MT (SMT), which rely on predictions drawn by comparing and analyzing segments in online bilingual corpora, and neural MT (NMT). Both SMT and NMT involve machine learning, but NMT leverages artificial neural networks to teach itself to accurately translate entire sentences, often as well as bilingual human translators (Wu et al., 2016). Although hybrid rule-based/SMT platforms such as SYSTRAN's Babel Fish were launched as websites in the late 1990s, it was not until the mid-2000s that a confluence of factors, including the proliferation of Wi-Fi networks in schools, the increased use of Internet-enabled devices by students, and—most significantly—the advent of Google Translate, sparked an acceleration in research into learner use of MT and its implications for language teaching and learning. Always evolving, in 2016 Google Translate implemented GNMT, its proprietary NMT model, which it continues to upgrade and whose accuracy improves with each user query.

In addition to a handful of dissertations and theses, empirical studies on the effects of MT use in L2 learning are still scarce. Therefore, and to recognize the contributions of a greater selection of sources, the scholarly works reviewed in this manuscript include two additional categories: argumentative essays and descriptions of classroom activities. These works can be organized into five interrelated strands, grouped by primary focus:

1. **MT and CAT Systems.** These publications explore the nature of MT systems and platforms, tracing their history and evolution, including the use of computer-assisted translation (CAT) tools by professional translators. They often focus on system capabilities, output quality or accuracy, and the evaluation of MT providers. Some of the most influential works in this strand are Aiken and Balan (2011), Aiken and Wong (2006), Ducar and Schocket (2018), Gaspari and Hutchins (2007), Groves and Mundt (2015), Hutchins (2010), Hutchins and Somers (1992), and Wu et al. (2016).

2. **MT in Translator Training.** These publications address the use of MT in translator training programs and translation courses, which are sometimes embedded in language programs. Relative to the distinction proposed by Klaudy (2003), these articles focus

on translation as a real-world professional skill rather than as a language-teaching tool. However, several of these works also address implications for L2 instruction and have influenced research in strand 3 below. Key works in this area include Allen (2003), Kenny and Way (2001), Kliffer (2005, 2008), La Torre (1999), Pym (2013), Sommers (2001, 2003), and Yang and Wang (2019).

3. **MT in Language Teaching and Learning.** This strand comprises essays, activity reports, and empirical studies that specifically address MT use in formal learning contexts. They discuss current and potential uses of MT to develop language skills (primarily translation and L2 writing), describe MT-supported activities, and report findings of experiments. Authors of these studies recommend best practices for using MT to support language learning. To date, the most influential works in this strand include Anderson (1995), Ball (1989), Belam (2003), Benda (2014), Corness (1986), Ducar and Schocket (2018), Enríquez Raído and Sánchez Torrón (2020), French (1991), Garcia (2010), Garcia and Cabot (2012), Garcia and Pena (2011), Giannetti (2016), Fredholm (2014, 2015a, 2015b, 2019), Jiménez-Crespo (2017), Lee (2020), Musk (2014), Niño (2004, 2008, 2020), O'Neill (2014, 2016, 2019b), Richmond (1994), Shei (2002a), Vold (2018), and Williams (2006).

4. **MT Use and Perceptions.** This strand consists of studies that report data on instructor and learner MT use, as well as perceptions, attitudes, and beliefs about MT tools in formal learning contexts. In addition to exploring frequency, motivations, and types of learner MT use, these studies often investigate instructor and learner beliefs about MT accuracy, usefulness, and academic integrity. Important studies of this type include Case (2015), Clifford et al. (2013), Eriksson (2021), Farzi (2016), Jin and Deifell (2013), Jolley and Maimone (2015), Knospe et al. (2019), Knowles (2016), Larson-Guenette (2013), Niño (2009, 2020), O'Neill (2019a), Shei (2002b), Stapleton (2005), and White and Heidrich (2013).

5. **MT as Academic Dishonesty.** Authors contributing works to this strand embrace the premise that MT use by learners in translator training and formal learning contexts equates to academic dishonesty or cheating. Thus, in addition to making the case for the MT-cheating connection, they focus on strategies for detecting, reacting to, and preventing unauthorized MT use. Representative examples of this strand are Correa (2011, 2014), Harris (2010), Innes (2019), Luton (2003), McCarthy (2004), Mundt and Groves (2016), O'Neill (2013), Somers et al. (2006), Stapleton and Ka Kin (2019), and Steding (2009).

In the following section, we map key emphases from studies pertaining to the latter three strands above onto a framework reflecting a series of questions commonly asked by both researchers and language practitioners: (1) How do language learners use MT tools? (2) What do instructors and learners think about MT tools? (3) How might MT use affect language learning? and (4) How should instructors respond to MT use by learners?

REVIEW OF THE LITERATURE

How Do Language Learners Use MT?

One of the primary strands of this research field relates to uses of and perceptions about MT, primarily those of learners. Most studies with this focus report empirical data collected through questionnaires or interviews (self-reports), although a few (Farzi, 2016; Fredholm, 2015a; Garcia & Pena, 2011; Knospe et al., 2019) also relied on classroom observations or screen recordings. In terms of learner use, these studies provide insights into how frequently learners use MT in instructed L2 contexts, the ways in which they tend to use it, and the reasons for such use.

Frequency of MT Use by Learners

There is ample evidence in the literature that students—by their own admission—are using MT with increasing frequency to support language learning, and especially on writing assignments. In an early survey of postsecondary Spanish learners in Australia, Pena (2011) found that 33% of students admitted to using MT on writing tasks. A couple years later, Clifford et al. (2013) reported that 71% of the Spanish learners at an American university used MT sometimes or often, with 88% reporting that they used MT for language learning purposes. In Larson-Guenette's (2013) study, 68% of German learners admitted using MT, and 69% of this percentage used it often. In Thailand, Sukkhwan (2014) surveyed L2 English postsecondary learners and found that 60% of them used MT more than twice a week. Fredholm (2015a) recorded secondary L1 Swedish students working on L3 Spanish writing activities and found that 63% of them used Google Translate to produce half or more of the L2 text.

Jolley and Maimone (2015) looked specifically at the use of Google Translate on writing assignments among intermediate L2 Spanish postsecondary learners. Results showed an astounding majority of participants (97.66%) had used MT at some point, with 74.11% reporting either frequent or occasional use. Similarly, Farzi (2016) found that all of the surveyed university ESL learners in his sample used MT at least some of the time, while 84% reported using it often or even more frequently. Surveying 310 university learners of either Spanish or French, O'Neill (2019a) found that 82.3% of them used MT sometimes, often, or always for non-graded work, a ratio that actually increased to 87.7% for graded work. High numbers of MT use were also reported in Knospe et al. (2019), who recorded and analyzed the behaviors of seven upper secondary L1 Swedish students engaged in L3 writing tasks in German. They found that all participants used online resources, including Google Translate, and that online searches accounted for a quarter of their total writing time. These findings, roughly covering the span of a decade, evidenced an already high use of MT in the early 2010s, which has become even more commonplace among most language learners today. This view, commonly held by language educators, is also borne out in observations made by Valijärvi and Tarsoly (2019) and Enríquez Raído and Sánchez Torrón (2020), who found that virtually all language students use MT for learning activities and that most of them use it frequently, especially on writing tasks.

How Learners Use MT

As with MT usage rates, studies conducted in the last decade reveal a strikingly consistent picture with respect to the types of assignments and specific purposes for which students use MT. In terms of task or assignment type, multiple surveys have found that students employ MT most frequently on L2 writing and writing-related tasks, such as pre-writing and revisions (Clifford et al., 2013; Jin & Deifell, 2013; Jolley & Maimone, 2015; Kol

et al., 2018; Larson-Guenette, 2013). For instance, Clifford et al.'s (2013) study showed that 43% of participants admitted to using MT on writing assignments, whereas in Jolley and Maimone (2015) and in Kol et al. (2018), students reported using MT at least occasionally in 85.16% and 83% of writing assignments, respectively.

In terms of segment length, research has found that students use MT most frequently to translate individual words or short phrases compared to paragraph-length or longer segments (Chandra & Yuyun, 2018; Clifford et al., 2013; Farzi, 2016; Jin & Deifell, 2013; Jolley & Maimone, 2015; Kol et al., 2018; Larson-Guenette, 2013). Jolley and Maimone's (2015) study exemplifies this trend, with 65.08% of participants reporting that they use MT for individual words, 14.57% for paragraph-length, and 11.72% for entire text. Farzi (2016) recorded students completing writing tasks and found their self-reported accounts to be accurate. An exception to this trend is found in Chen (2020), who reported that 69% of her participants used Google Translate to translate paragraphs or entire texts. Finally, in a less explored area of MT use, Stapleton (2005) and Wuttikrikunlaya et al. (2018) investigated the effects of proficiency level on MT use. Both of their studies found that low-proficiency learners tended to translate longer segments, such as entire sentences, while more advanced leaners reported mainly using MT to check the meaning of individual words. Although most self-reported MT use by students mimics traditional dictionary usage, the fact that many learners report translating paragraphs or entire texts when engaged in L2 writing assignments is disconcerting.

Reasons for Learner MT Use

The fact that language learners use MT frequently is evidence that their skepticism regarding MT accuracy and their ambivalence about the effectiveness and ethicality of MT use are often overridden by the rationales they cite for using it. When asked why they turn to MT so frequently, learners mentioned the fact that MT tools are free and easily accessible (Jin & Deifell, 2013; Niño, 2020; Sukkhwan, 2014) and that they are fast and convenient timesavers (Clifford et al., 2013; Fredholm, 2015a; Jin & Deifell, 2013; Larson-Guenette, 2013; Xu & Wang, 2011). Looking at the issue from a psycholinguistic perspective, Selcuk et al. (2019) suggested that low learner confidence and high levels of anxiety related to L2 writing may be other motivating factors to use MT. However, to the best of our knowledge, the role of affective and cognitive factors among learners and their MT use has not yet been explored.

What Do Instructors and Learners Think about MT Tools?

Instructor Perceptions about MT

To understand the role instructors expect MT tools to play in the L2 classroom, researchers have examined their perceptions, attitudes, and beliefs regarding MT use (Baker, 2013; Case, 2015; Clifford et al., 2013; Eriksson, 2021; Farzi, 2016; Jin & Deifell, 2013; Jolley & Maimone, 2015; Knowles, 2016; Niño, 2009, 2020; O'Neill, 2019a; White & Heidrich, 2013). Their findings show various degrees of knowledge about MT capabilities and acceptance of MT use by students. For example, while many instructors disapproved of MT use by their students (Clifford et al., 2013; Eriksson, 2021; Knowles, 2016) or believed that MT use or reliance had a negative effect on language learning (Baker, 2013; Stapleton & Ka Kin, 2019), results also showed that some instructors had positive impressions of MT and believed it had a place in L2 instruction (Clifford et al., 2013; Eriksson, 2021; Knowles, 2016; Niño, 2009;

Stapleton & Ka Kin, 2019). Instructors also consistently voiced concern over MT output quality or accuracy (Niño, 2009; Jolley & Maimone, 2015; Stapleton & Ka Kin, 2019).

Results also show that instructors' perceptions and beliefs about MT use are not always straightforward. Instructors often hold nuanced, conflicting, and even contradictory views regarding the place and value of MT in language instruction, as shown by Baker (2013), Case (2015), and Eriksson (2021). Instructors' conflicting views often depend on how MT is used, the frequency of use, and the types of assignments involved. Despite the general skepticism and concerns about how much MT use facilitates learning, many instructors believed that student MT use is unavoidable, and that instructors and students alike would benefit from training regarding how to best use MT tools to support language teaching and learning (Baker, 2013; Benda, 2014; Case, 2015; Eriksson, 2021; Jolley & Maimone, 2015; Stapleton & Ka Kin, 2019).

Learner Perceptions about MT

Learner perceptions, attitudes, and beliefs about MT and MT use in translation and L2 writing tasks have been investigated in several studies (Chen, 2020; Clifford et al., 2013; Enríquez Raído & Sánchez Torrón, 2020; Farzi, 2016; Fredholm, 2015a; Garcia & Pena, 2011; Giannetti, 2016; Jin & Deifell, 2013; Jolley & Maimone, 2015; 2011 Larson-Guenette, 2013; Lee, 2020; Niño, 2009, 2020; O'Neill, 2019a; Shei, 2002b, Sukkhwan, 2014; Tsai, 2019; Valijärvi & Tarsoly, 2019; White & Heidrich, 2013; Xu & Wang, 2011). In Niño (2009), 81% of learners who completed post-editing activities using MT reported believing that MT helped them learn the language, while 75% noted that MT post-editing boosted their confidence. Clifford et al. (2013) found that 63% of learners thought MT was useful for language learning sometimes, and 31% believed it to be always useful. In Jolley and Maimone (2015), 55.5% of learners agreed or strongly agreed that MT had a positive impact on language learning, while in O'Neill (2019a), 75.6% of participants expressed positive opinions of MT use in general.

Similarly, 88% of participants in Lee's (2020) study felt that using MT as a revision tool helped improve their writing skills. Enríquez Raído and Sánchez Torrón (2020) reported comparable findings. L1 Chinese ESL learners in Tsai's (2019) study (N=124) completed a 5-point Likert scale questionnaire about their perceptions of MT in L2 writing. An average of 3.8 learners claimed they benefited from using MT for lexical items, 3.7 believed it enhanced final English renderings, and 3.52 we satisfied with MT outputs overall. Niño (2020) looked at students' perceptions of both written translations and oral output produced by online MT tools, including Google Translate. She found that over 60% of learners found MT written translation to be useful or very useful for different tasks, while 67% felt using MT helped improve their writing.

But learner perceptions of MT are not always positive. In Niño (2020), students showed a different degree of acceptability regarding audio output, which presented issues of comprehensibility, with 40% reporting not finding results useful. They also reported believing that MT accuracy varied by language pair. In Clifford et al. (2013), 78% of the participants judged MT output as somewhat accurate at best. Similarly, 71.1% of the learners surveyed by Jolley and Maimone (2015) considered Google Translate output to be somewhat accurate, with the translation of longer segments perceived as less reliable. On the other hand, White and Heidrich (2013) noted that, despite trusting the accuracy of online MT tools, students reported not being able to express their own voice or not noticing improvements in their writing style.

Learners have also been asked whether they believed that using MT should be considered cheating. In White and Heidrich's (2013) study, learners expressed the sense they had cheated by using MT tools. In interviews with learners, Baker (2013) reported that students felt conflicted about using MT and that the awareness of plagiarism concerns made them reluctant to justify such use. In Niño (2020), 90.9% of participants said they would object to policies banning online MT tools use because they did not believe they induced plagiarism. This nuanced perception of cheating appears in many studies, which also show that learners are capable of making distinctions between acceptable and unacceptable uses of MT in the L2 classroom. For example, while learner perceptions about MT were overall positive in Jolley and Maimone's (2015) study, with 74.80% reporting that the use of Google Translate on writing assignments was somewhat or completely ethical, 86.72% of learners thought that whether using MT equated to cheating or not depended on how it was used (e.g., type of assignment, segment length, etc.).

Overall, learner use and perception studies reveal that students hold mixed views regarding the suitability, reliability, and ethicality of MT tools. While results suggest that learners generally hold positive views of MT, believing that it has the potential to support their learning and improve the quality of L2 writing, these positive perceptions are tempered by concerns regarding MT accuracy, awareness of its shortcomings, and conflicted views of what constitutes ethical behavior (Baker, 2013; Jin & Deifell, 2013; Jolley & Maimone, 2015; Larson-Guenette, 2013; Niño, 2020; Valijärvi & Tarsoly, 2019; White & Heidrich, 2013). Additional issues relating to academic dishonesty, including instructor views, are discussed in greater length below.

How Might MT Affect Language Learning?

Using MT to Raise L2 Awareness

Many early publications that address the suitability of MT for language learning describe MT translation activities involving error identification and correction, pre-editing, and post-editing as a means of raising learner awareness and thus enhancing their explicit knowledge of L2 vocabulary and grammar. Pioneers of this approach include Ball (1989), Corness (1986), French (1991), Richmond (1994), Shei (2002a), and Somers (2001, 2003), who famously referred to strategies that involve repair of faulty MT output as "using the MT as a bad model" (Somers, 2003, p. 327). Richmond (1994), the first researcher to make theoretical connections between MT use and second language acquisition (SLA) theory, proposed the design of back-translation and pre-editing activities to promote crosslinguistic comparisons and enhance learners' awareness of L2 grammaticality (p. 66). He suggested that such activities increased attention to form and maximized input processing. Similarly, Belam (2003) argued that post-editing requires learners to analyze both the source and the target text in detail, which increases vocabulary and syntactic gains.

More recently, Niño (2008) evaluated the accuracy of MT output that had been post-edited by advanced learners compared to non-MT-assisted output. She noted that having learners critically evaluate raw MT output creates opportunities for error awareness and correction and negotiation of meaning. Many studies that followed have suggested similar classroom activities involving MT output manipulation, also based on the perceived benefits of increased learner metalinguistic awareness (Benda, 2014; Correa, 2014; Enkin & Mejías-Bikandi, 2016; Valijärvi & Tarsoly, 2019). Lee's (2020) study offers some evidence of these

benefits. Looking at the quality of L2 writing samples produced by postsecondary EFL learners, she found that those who compared their own translations with MT output produced better texts, a result she attributed to increased attention to form and metalinguistic awareness promoted by the pedagogical use of MT.

The Impact of MT on L2 Writing Quality

The first empirical study to explore the potential benefits of MT use for L2 writing was Niño (2004). She conducted two experiments involving 13 advanced learners of L2 Spanish. In the first experiment, students post-edited faulty MT output with the aid of the English source text and reference materials. She found that the post-edited texts presented slightly higher percentages of errors in grammar, but almost identical numbers of spelling and lexical errors compared to raw MT output. Post-edited texts also showed a wider range of grammatical and lexical errors. In the second experiment, she compared post-edited MT output to translations done by L2 learners without the use of MT. Results showed that post-edited texts had fewer grammatical errors than the L2 learner translations but slightly higher lexical and spelling errors. Interestingly, the only category in Niño's study in which either learner translation or post-editing generated better results than raw MT output was vocabulary.

Other studies examined specific aspects of written texts, such as lexical diversity, complexity, and grammatical control, and present a more favorable outlook of MT use for pedagogical purposes. For example, in Garcia and Pena (2011), L2 Spanish learners completed writing tasks with or without the assistance of the Tradukka interface. Final drafts were analyzed in terms of the writing process, total number of words, and quality of output. Their findings showed that MT use allowed beginner learners to write more and with less effort. Fredholm (2014) investigated the L3 Spanish written production of two groups of L1 Swedish students—One group with full access to online resources and spell-checking tools, and a control group with no access to either resource. In one semester, students completed four compositions, which were measured for morphosyntactic and lexical-pragmatic accuracy. Comparing both groups, he found that the group with access to online MT tools produced significantly fewer errors in mood selection, noun agreement, lexical choices, and syntax, with a higher ratio of accurate sentences per essay. While the quality of writing output for the MT group was higher, what the study does not make clear is whether the use of MT helped students make fewer mistakes over time. Using the same research design, Fredholm (2015b) further investigates differences in fluency, and lexical and grammatical complexity, finding similar results overall to his 2014 study. Along the same lines, Farzi (2016) observed the writing behavior of university ESL learners over the course of 12 weeks and found that approximately 67% of the time learners' writing showed increased variety of vocabulary and grammatical structures when they employed MT tools.

Some of these studies looked specifically at the use of Google Translate. In one of them, Giannetti (2016) reported that Google Translate helped improved students' abilities to construct grammatically correct sentences and reduced the number of semantic errors in L2 Spanish writing assignments. Kol et al. (2018) observed that students wrote longer texts and attained richer vocabulary profiles and readability levels when using Google Translate. Fredholm (2019) found that L1 Swedish students using Google Translate produced L3 Spanish compositions with significantly higher lexical diversity most of the time, while Tsai (2019) reported that Google Translate-generated texts received higher scores on measures of grammar and vocabulary. Finally, Lee (2020) had L1 Korean students learning EFL compare

their own translations to MT output (Google Translate or Papago Web Translator) to facilitate error correction and found that they significantly improved the quality of their final drafts.

O'Neill (2014, 2016) also looked at differences in the quality of specific writing components but introduced MT training as new variable. He examined L2 French compositions produced by L1 English postsecondary students under three conditions: no MT use, MT use after MT training, and MT use with no training. In his studies, samples from at least one task were rated significantly higher for both the MT training and no training groups when compared to the no MT use group on grammar, content, comprehensibility, and spelling. In addition to specific components, O'Neill (2014, 2016) investigated the effects of MT on global ratings of L2 French compositions, finding that both uses of MT resulted in overall improved writing quality. In a more recent study, O'Neill (2019b) analyzed over 1,000 French and Spanish compositions written by L1 English speakers with the help of Google Translate, an online dictionary, or no tools. Results again showed an advantage for the use of MT over no tools and online dictionaries. Taken together, O'Neill's findings point to improved writing quality when learners use MT, especially when MT training is provided.

The combined findings of the studies above suggest that MT tools do indeed belong in the L2 classroom as a resource for translation and writing tasks. They provide evidence that MT facilitates the production of longer and more accurate texts, which tend to be graded more favorably. Nonetheless, whether such activities result in language learning (or acquisition) is not clear. When discussing L2 learning, most of these studies measure success through the quality of the writing samples but do not necessarily assess the development of writing skills or gains in grammatical and lexical knowledge. In fact, researchers have stressed that the demonstrated gains in writing quality that result from MT assistance are not likely permanent and that overreliance on MT may actually hinder L2 learning and writing development (Fredholm, 2015a, 2019; Garcia and Pena, 2011; Giannetti, 2016; O'Neill 2016, 2019b; Valijärvi & Tarsoly, 2019).

MT and Language Development

The studies described above focus primarily on the benefits of MT use for L2 writing performance. A common thread in these studies is the belief that writing and translation activities involving different degrees of MT assistance draw learners' attention to form and increase their L2 metalinguistic awareness, helping develop knowledge of the L2. Indeed, the cognitive processes promoted by such activities are well documented in the SLA literature as facilitating L2 development. These include focused attention (Tomlin & Villa, 1994), noticing and awareness (Leow, Johnson, & Zárate-Sández, 2011; Robinson, 1995; Schmidt, 1990), input processing (McLaughlin, 1990; VanPatten, 2004), and depth of processing (Hulstijn & Laufer, 2001; Leow, 2015; Leow & Mercer, 2014). Similarly, the actions that promote these kinds of processes, such as solving communication breakdowns, negotiating meaning, detecting and correcting errors, and comparing crosslinguistic contrasts and similarities are also frequently discussed in SLA research. However, connections between language acquisition theory, learning outcomes, and MT use are rarely found in the MT literature. As noted above, one exception is Richmond (1994), who explores the nature of language representation and models of input processing, proposing the design of activities centered on the writing process and on language analysis rather than on the writing output.

Most importantly, very few empirical MT studies have sought to investigate actual language gains promoted by different MT tools and activities. While one can argue that writing and translation products offer a window into the learners' implicit language knowledge

(though their grammaticality judgments and choices), the evaluation of written samples in MT studies mostly reflects the learners' writing strategies and translation skills. Text samples written with the help of MT tools do not necessarily capture what learners know, since learners do not rely solely on their own resources for language production. It is also not possible to infer from these samples if new language knowledge has been acquired.

One of the few researchers to assess L2 development was O'Neill (2014, 2016, 2019b). In O'Neill (2014), results showed significant differences in composition scores between groups who used and did not use MT. After receiving MT training, one of the groups received significantly higher scores in a second composition. However, as noted by O'Neill himself, the higher scores on the second composition could have resulted from more effective use of MT due to training rather than learners' increased language knowledge. O'Neill (2016) addressed this issue by administering pretests and posttests in the form of compositions without access to spell-checkers or online MT, thus eliminating the conflating variable. The pretests and posttests were scored for comprehensibility, content, vocabulary, spelling, and syntax, and no differences were found between groups before or after training, during which some groups used MT in L2 writing tasks. A reading pretest was used as a measure of proficiency, but it was not administered as a posttest to assess proficiency gains.

In O'Neill (2019b), L2 French and L2 Spanish compositions were used as pretests, posttests, and delayed posttests, which were assessed for fluency, accuracy, and complexity. This time, O'Neill investigated use or not of MT, exposure or not to MT training, and two different online tools, Google Translate and WordReference (an online dictionary), with a total of five experimental groups. He found that when the groups who received training in either online tool and were allowed to use them to complete the two experimental writing tasks, they significantly outperformed the control group and the groups using the tools without having received training. However, the results for the posttest showed that when the MT-assistance was removed, the groups who received MT training actually performed significantly worse than the other groups, though no significant differences were observed in the delayed posttest. O'Neill's (2019b) study, therefore, paints a positive view of the use of MT for improving the quality of L2 writing output but a relatively negative view of the use of MT for the development of L2 writing proficiency.

Perhaps the only other researcher to assess learning is Fredholm (2019). He analyzed the lexical diversity of L3 Spanish compositions written with and without MT assistance and had his participants complete pretests and posttests (essay and grammar test components) without access to MT tools. No significant differences between pretest and posttest essays across groups were found. The results for the grammar pretests and posttests were not reported. One observation to be made is that while the studies above found no positive gains in language development as an effect of MT use, the only measure used to assess linguistic knowledge were writing samples. Clearly, research into connections between MT use and any language development is still in the early stages. But it is also possible that potential linguistic gains were not captured by the complexity of the writing process or that the type and frequency of MT use during the experiments were insufficient to yield measurable effects.

How Should Instructors Respond to MT Use by Students?

The well-documented reality that most language learners use MT frequently in writing tasks confronts language instructors with a series of choices. Recommendations for responding to this situation tend to reflect one of two mindsets or assumptions about MT use: It is either

considered cheating or seen as a possible resource to facilitate learning (White & Heidrich, 2013). Researchers who consider learner MT use a form of academic dishonesty often focus on its drawbacks and recommend strategies for detection, reaction, and prevention, whereas those who see MT as a resource generally underscore its benefits and suggest integrating MT tools into the language curriculum.

Instructor Beliefs about MT and Academic Dishonesty

Many influential essays on MT use describe it as problematic, unethical, and detrimental to meaningful learning (Correa, 2011, 2014; Harris, 2010; Luton, 2003; McCarthy, 2004; Mundt & Groves, 2016; Somers et al., 2006; Steding, 2009). Luton (2003) is the first researcher to explicitly connect learner MT use to academic dishonesty. Highly skeptical of MT capabilities, she states that MT allows students to avoid engaging with assignments and learning how to write in the target language. Despite these concerns, she noted that guided MT activities may benefit advanced learners. For McCarthy (2004), MT-produced submissions passed off as student work are unacceptable. He points out that work produced with no intellectual input from the student has no instructional value, wastes the instructor's time, and defeats the purpose of ongoing assessment. Concerned that MT reliance interferes with meaningful target language use and disrupts the process approach to L2 writing, Harris (2010) claims that many teachers believe learner MT use to constitute a form of cheating and that accepting MT-assisted texts is unfair to students who complete writing tasks on their own. Similarly, both Correa (2011) and Steding (2009) cite the use of online translators as a major academic dishonesty issue and assert that undisclosed use of MT on assignments should be considered cheating and treated accordingly.

Even though many instructors believe that MT use constitutes cheating, survey data suggest that this view is far from absolute. For example, in Clifford et al. (2013), 42% of instructors stated that MT use was cheating and 37% said that whether MT equated to cheating depended on how students used it. Jolley and Maimone (2015) found that 12.82% of the instructors always considered MT use cheating, while 82.50% also said it depended on how students used it. A similar scenario was reported by Case (2015). She observed that, although 22 of the 35 instructors agreed that MT was cheating, they sometimes expressed contradictory sentiments, and the extent of agreement depended on factors such as segment length, context, assignment, type, and learner level. In one of the first studies to ask language instructors about MT use and academic dishonesty, Correa (2011) noted that there was no consensus among instructors, and that MT use was considered less serious than other forms of cheating. Baker (2013) observed that instructors felt conflicted about MT technology because of fears of plagiarism and potential negative effects on language learning, and Knowles (2016) found that 35% of the instructors equated MT use with cheating, a percentage that held steady even after completion of an MT familiarity module.

Signs of MT Use

Several researchers have addressed the telltale signs indicative of student MT use, often equating them with perceived deficiencies in MT output. In his evaluation of a Hebrew-to-English MT system, Anderson (1995) found that its most common errors included problems with homographs, mistranslation of prepositions, untranslated words, incorrect translation of verb tenses, failure to recognize proper nouns, and incorrect negation. For Japanese-to-English translation, Innes (2019) cited erroneous or inappropriate passive voice use and

preposition mistakes as red flags that helped raters correctly identify MT use. Luton (2003) noted that MT systems struggled with idiomatic expressions, often left words in the source language, and translated proper nouns, contrary to usual translation practice. Steding (2009) grouped common MT mistakes into four categories (spelling, vocabulary choice, grammar, content and style) and mentioned untranslated and misspelled words, incorrect words, unexpected words, and strange/odd words as specific examples. Compiling observations from other studies, Correa (2014) put together a more comprehensive list of MT use indicators, citing literal translations, grammatical inaccuracies, inability to account for cultural references and other extralinguistic or contextual issues, "unnatural" writing, misspelled words, difficulty with idioms, errors that humans do not make, and the translation of proper nouns.

Though highly critical of student MT use, Steding (2009) noted that constant improvements in online MT system capabilities mean that instructors interested in detecting such use should also consider what these systems do well. In this respect, his list of telltale signs of MT use also emphasized the absence of expected mistakes, consistently correct grammar forms, and unexpected levels of cohesion/organization in the second language. In the years since Steding's essay, and particularly since Google Translate's upgrade to an NMT model in 2016, there has been a shift in detection criteria from weaknesses or mistakes to signs of uncharacteristically good production. Ducar and Schocket (2018) summarize this shift and note some specific red flags for instructors:

> As MT technologies continue to improve, identifying translation "mistakes" will likely become increasingly difficult for language instructors. Instead, it will be the technology's subtle successes, rather than its breakdowns, that will signal MT use. Telltale signs of MT use by students include producing verb tenses that have not yet been studied, using excessive advanced vocabulary, and producing subordinated complex clauses, and even the absence of prepositional errors that are typically produced by lower- and intermediate-level learners. (p. 787)

Instructor Ability to Detect MT Use

Despite the attention focused on signs of MT use, just three studies to date (Innes, 2019; O'Neill, 2013; Stapleton & Ka Kin, 2019) have investigated the extent to which human raters are able to reliably detect the use of MT by language learners. O'Neill (2013) reported that raters reviewing writing tasks produced by 32 university L2 French learners were able to judge correctly whether MT was used with a 70.70% accuracy rate. However, he admits that this fact alone does not support the conclusion that MT-assisted compositions can always be easily identified. In Innes' (2019) study, five advanced L2 learners of English (native speakers of Japanese) translated articles into English. Those translations were then paired with Google Translate-generated versions and presented to 17 native-speaking English teachers in Japan, who were asked to identify which text in each set was MT-generated. The percentage of raters making the correct identification varied by text set from 58.82% to 93.75%, with a mean score of 74.04%. Stapleton and Ka Kin (2019) had raters grade a random mixture of Google Translate-generated and unassisted essays. The Google Translate-generated essays received higher scores, but results were not statistically significant, leading the researchers to conclude that the raters were not able to recognize Google Translate use, a finding they attributed to its improved quality.

MT as Cheating and the Detect-React-Prevent Response

The recommendations that follow all center on three main processes for responding to unauthorized student MT use: detect, react, and prevent. They begin with McCarthy (2004), who shared 12 solutions gleaned from input solicited from translation training students, as well as his own recommendations. These solutions included having course objectives that addressed MT use and abuse, raising students' awareness of their own abilities and advantages relative to MT, and designing translation and writing assignments and assessments with features resistant to MT use. Similarly, Steding's (2009) influential approach holds that language educators encountering unauthorized MT use have a threefold task: the detection of students who give in to temptation, the choice of an adequate reaction, and the implementation of preventive measures. According to him, the keys to reliable detection are for instructors to know their students and the kinds of mistakes they tend to make, as well as the capabilities and limitations of MT technologies. In terms of reaction, he advises that the offense be treated as a violation of department and course policies when charges of academic dishonesty are substantiated. Regarding prevention, Steding recommends having a clear syllabus policy that addresses MT use, showing students examples of bad MT output, requiring that learners submit statements of authorship, and creating so-called smart assignments for which MT use would not be feasible. Harris (2010) devised his own version of a warning protocol, pressing instructors to educate learners with awareness-raising activities, confront them with suspicions of cheating, and make them re-do assignments when cheating is suspected. Correa (2011) also suggests that instructors follow a multistep model to prevent unauthorized MT use, recommending that language instructors work collaboratively to establish and enforce academic dishonesty policies. In a subsequent essay, Correa (2014) proposes a modified plagiarism response with a greater focus on prevention, taking the view that prevention is more efficient and less costly than punishment. While not rejecting the importance of detection and enforcement for committed cheaters, she suggests having learners engage with MT output to raise metalinguistic awareness in hopes that such activities will discourage them to cheat.

MT as Resource and the Integrate-Educate-Model Approach

Not all researchers subscribe to the view that MT use should be discouraged. Many, on the contrary, recommend that instructors focus on the pedagogical implications of MT use, integrating it into the curriculum and modeling best practices for language learning. Recognizing that translation technologies are here to stay, Jiménez-Crespo (2017) suggests that instructors accept the reality of digital natives, considering the didactic role of MT in the 21st century classroom. Echoing this sentiment, Vold (2018) emphasizes that researchers and language teachers need to shift their attention to the ways in which MT can support language learning, not just communication. Commenting on possible benefits for students trained in MT use, O'Neill (2019b) concludes that since there is no failproof method to prevent MT use outside the classroom, providing MT training that shows learners how to utilize MT more appropriately and responsibly may be more useful. Many other studies published in the past six years (e.g., Ducar & Schocket, 2018; Farzi, 2016; Fredholm, 2015a, 2019; Giannetti, 2016; Jolley & Maimone, 2015; Knospe et al., 2019; Knowles, 2016; Mundt & Groves, 2016; Musk, 2014; Niño, 2020; O'Neill, 2016, 2019a; Stapleton & Ka Kin, 2019) support the argument for greater integration of MT into the language curriculum, including clearer definitions of appropriate and inappropriate uses and purposeful training.

This recent movement away from the detect-react-prevent mindset and toward the integrate-educate-model approach appears to have been driven in large part by the realization that, given the ease with which students can access MT, formally prohibiting MT use is fruitless (Fredholm, 2019, Conclusions and Discussion). Framing these conditions as an opportunity, Ducar and Schocket (2018) make the point that the issue is no longer whether instructors can prevent learners from consulting MT technologies but rather how to help them understand that progressing toward greater proficiency and using MT tools ethically are critical 21st century skills. Stapleton and Ka Kin (2019) make a more urgent appeal for acknowledging the importance of MT tools as strategies for reading and writing in a foreign language. Citing multiple benefits of MT for 21st century language learners, Niño's (2020) advice to instructors is not to fear MT and its potential role in language learning, let alone ban or ignore it. Instead, she argues, they should seek opportunities to integrate MT into the L2 classroom in meaningful and realistic ways.

CONCLUSIONS

Summary and Key Takeaways

It is clear from the foregoing review that research exploring connections between MT and language teaching and learning constitutes a well-established and rapidly expanding field. In just over 30 years, this domain has produced important insights with implications for language education. In terms of student MT use and perceptions, it is clear that a majority of learners use MT on writing tasks very frequently, primarily to look up words and phrases. Learners are largely aware of MT's limitations but nonetheless see it as a convenient and beneficial aid to their language learning efforts. Both instructors and learners hold nuanced and often conflicting views on the suitability of MT in the L2 classroom and, although many instructors object to student use of MT, some see a pedagogical role for it, supporting additional MT training for themselves and their students. Overall, the literature examining the potential benefits of MT for language learning has not yielded conclusive results, but points toward two directions: MT may be a useful tool for enhancing learners' metalinguistic knowledge, and it helps students perform better on translation and L2 writing tasks.

Strategies for responding to the inevitable use of MT by students align with the instructors' views regarding the ethicality of its use. The research clearly indicates that many instructors consider unauthorized MT use a form of academic dishonesty, leading many scholars to recommend strategies rooted in the detect-react-prevent approach, even as support builds for the integrate-educate model. Signs of MT use in the literature include both shortcomings of MT tools and areas in which they are effective. However, although instructors are usually aware of these signs, their ability to reliably detect such use has not been proved conclusively. Finally, researchers in the past few years have increasingly characterized MT use by students as inevitable and possibly beneficial, supporting the idea of exploring ways to use it for pedagogical purposes.

Directions for Future Research

To date, MT studies have explored important questions for language education but have only just begun to produce empirical data to support researchers' and instructors' intuitions and

observations, especially if we consider the methodological rigor required today in the SLA field. Future research should engage current methodologies in SLA and cognitive linguistics, which include the use of language background surveys, assessment of participants' L2 proficiency, and reports that present detailed information about participants, treatment materials, data collection instruments, and coding procedures. Furthermore, studies focused on the impacts of MT use on language teaching and learning should make explicit connections between pedagogical practices and SLA theories, showing an understanding of the cognitive processes that promote language acquisition and the concrete ways in which proposed tasks involving MT facilitate these processes.

Much of what is known so far about learner use and perceptions of MT tools applies to Spanish, German, English, and French as second or third languages. More research is needed, therefore, to understand the role of MT tools when other language pairs are involved and when the learning contexts and traditions differ significantly from those that have been already studied. A closer look at the effects of L2 proficiency, learner individual differences, and teaching methods on MT use and effectiveness may also contribute valuable information for instructors interested in integrating MT into the language curriculum.

The implications of MT training are another important aspect that has just begun to be explored. More research is needed to help instructors identify exactly what training of this type should entail to best assist them in making decisions about materials design, course policies, and grading and how such training might help learners make decisions about their learning process. In terms of MT use detection, studies have been able to identify signs of MT use and how well instructors can distinguish between texts produced with or without MT assistance. The next step seems to be the identification of actual text characteristics in writing samples that have been correctly identified as MT-assisted or non-MT-assisted.

Finally, advances in the field of MT use for L2 learning and teaching will require research that looks more closely at language and communicative competence development. To capture different types of knowledge and subtle changes, a variety of instruments should be used to measure discrete aspects of language or language proficiency. Research is needed on both short-term and longitudinal effects of MT use on language processing, memory, and information retrieval. For example, instead of relying uniquely on writing performance as linguistic data, researchers can assess intake, uptake, and restructuring mechanisms, as well as online levels of attention and noticing. Also, as we have seen, most researchers assert that MT use has the potential to raise learners' awareness of language and crosslinguistic contrasts. Therefore, future research should prioritize measuring gains in explicit knowledge and improvements in learners' grammatical sensitivity and analytical skills.

NOTES

[1]Since most of the literature reviewed herein focuses on second language (L2) settings and rarely addresses participants' language backgrounds in detail, we use the abbreviation L2 in a generic sense to refer to the teaching and learning of additional languages except when referencing studies that specify that participants are third language (L3) learners.

[2]See Németh (2019) for an accessible, accurate overview of this topic. For more detailed accounts, see Garg and Agarwal (2019), Hutchins (2010), and Wu et al. (2016).

REFERENCES

Aiken, M., & Balan, S. (2011). An analysis of Google Translate accuracy. *Translation Journal*, *16*(2).

Aiken, M., & Wong, Z. (2006). Spanish-to-English translation using the Web. *Southwest Decision Sciences Institute.*

Allen, J. (2003). Post-editing. In H. Somers (Ed.), *Computers and translation. A translator's guide, Vol. 35* (pp. 297-317). John Benjamins.

Anderson, D. D. (1995). Machine translation as a tool in second language learning. *CALICO Journal*, *13*(1), 68-97.

Atkinson, D. (1993). Teaching in the target language: A problem in the current orthodoxy. *The Language Learning Journal*, *8*(1), 2-5.

Baker, C. (2013). *Student and instructor perceptions of online translators in English composition* [Master's thesis, Mississippi State University]. ProQuest Dissertations Publishing.

Ball, R. V. (1989). Computer-assisted translation and the modern languages curriculum. *The CTISS File*, 8, 52-55.

Belam, J. (2003, September 23-27). Buying up to falling down: A deductive approach to teaching post-editing [Paper presentation]. MT Summit IX Workshop on Teaching Translation Technologies and Tools, New Orleans, LA, United States.

Benda, J. (2014). Google Translate in the EFL classroom: Taboo or teaching tool? *Writing & Pedagogy*, *5*(2), 317-332.

Case, M. (2015). Machine translation and the disruption of foreign language learning activities. *eLearning Papers*, 45, 4-16.

Chandra, S. O., Yuyun, I. (2018). The use of Google Translate in INEFL essay writing. *LLT Journal*, *21*(2), 228-238.

Chen, W. (2020). Using Google Translate in an authentic translation task: The process, refinement efforts, and students' perceptions. *Current Trends in Translation Teaching and Learning E*, 7, 213-238.

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.

Cook, G. (2010). *Translation in language teaching: An argument for reassessment*. Oxford University Press.

Cook, V. (2001). Using the first language in the classroom. *The Canadian Modern Language Review*, *57*(3), 402-423.

Corness, P. (1986). The ALPS computer-assisted translation system in an academic environment. *Translating and the Computer*, 7, 118-127.

Correa, M. (2011). Academic dishonesty in the second language classroom: Instructors' perspectives. *Modern Journal of Language Teaching Methods*, 1, 65-79.

Correa, M. (2014). Leaving the "peer" out of peer-editing: Online translators as a pedagogical tool in the Spanish as a second language classroom. *Latin American Journal of Content and Language Integrated Learning (LACLIL)*, *7*(1), 1-20.

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.

Enkin, E., & Mejías-Bikandi, E. (2016). Using online translators in the second language classroom: Ideas for advanced-level Spanish. *Latin American Journal of Content and Language Integrated Learning (LACLIL), 9*(1), 138-158.

Enríquez Raído, V., & Sánchez Torrón, M. (2020) Machine translation, language learning and the 'knowledge economy.' In M. Filimowicz & V. Tzankova (Eds.), *Reimagining communication: Action* (pp.155-171). Taylor and Francis/Routledge.

Eriksson, N. L. (2021). *Google Translate in English-language learning: A study of teachers' beliefs and practices* [Unpublished master's thesis]. Dalarna University.

Farzi, R. (2016). *Taming translation technology for L2 writing: Documenting the use of free online translation tools by ESL Students in a writing course* [Unpublished doctoral dissertation]. University of Ottawa.

Fredholm, K. (2014). Effects of online translation on morphosyntactic and lexical-pragmatic accuracy in essay writing in Spanish as a foreign language. In S. Jager, L. Bradley, E. J. Meima, & S. Thouësny (Eds.), *CALL Design: Principles and Practice. Proceedings of the 2014 EUROCALL Conference* (pp. 96-101), Groningen, The Netherlands.

Fredholm, K. (2015a). El uso de traducción automática y otras estrategias de escritura digital en español como lengua extranjera. *Estudios de Lingüística Aplicada*, 62, 9-32.

Fredholm, K. (2015b). Online translation use in Spanish as a foreign language essay writing: Effects on fluency, complexity and accuracy. *Revista Nebrija de Lingüística Aplicada a la Enseñanza de las Lenguas*, (18), 1-18.

Fredholm, K. (2019). Effects of Google translate on lexical diversity: Vocabulary development among learners of Spanish as a foreign language. *Revista Nebrija de Lingüística Aplicada a la Enseñanza de las Lenguas*, *13*(26), 98-117.

French, R. J. (1991). Machine translation. In W. Brierley & I. R. Kemble (Eds.), *Computers as a tool in language learning* (pp.55-70). Ellis Horwood Limited.

García, I. (2010). Can machine translation help the language learner*? ICT for Language Learning,* 3rd edition. PIXEL.
García, I., & Cabot, C. (2012). *Does machine translation support language learning?* In *EDULEARN12 Proceedings* (pp. 4511-4519). 4th International Conference on Education and New Learning Technologies, Barcelona, Spain.
Garcia, I., &, Pena, M. I. (2011). Machine translation-assisted language learning: Writing for beginners. *Computer Assisted Language Learning*, *24*(5), 471-487.
Garg, A., & Agarwal, M. (2019). Machine translation: A Literature review. *ArXiv, abs*/1901.01122.
Gaspari, F., & Hutchins, J. (2007). Online and free! Ten years of online machine translation: Origins, developments, current use and future prospects. In B. Maegaard (Ed.), *MT Summit XI–Proceedings* (pp. 199-206)*,* European Association of Machine Translation, Copenhagen, Denmark.
Giannetti, T. R. (2016). *Google Translate as a resource for writing* [Unpublished master's thesis]. St. John Fisher College.
Groves, M., & Mundt, K. (2015). Friend or foe? Google Translate in language for academic purposes. *English for Specific Purposes*, 37, 112-121.
Harris, H. (2010). Machine translations revisited: Issues and treatment protocol. *The Language Teacher*, *34*(3).
Hulstijn, J., & Laufer, B. (2001). Some empirical evidence for the Involvement Load Hypothesis in vocabulary acquisition. *Language Learning*, *51*(3), 539-558.
Hutchins, W. J. (2010). Machine translation: A concise history. *Journal of Translation Studies*, *13*(1-2), 29-70.
Hutchins, W. J., & Somers, H. (1992). *An introduction to machine translation.* Academic Press.
Innes, A. R. B. (2019). Differentiating between translation and student translation: Red flags salient lexicogrammatical features. *Lublin Studies in Modern Languages and Literature*, *43*(4), 1-13.
Jiménez-Crespo, M. A. (2017). The role of translation technologies in Spanish language learning. *Journal of Spanish Language Teaching, 4*(2), 181-193.
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.
Kenny, D., & Way, A. (2001). Teaching machine translation and translation technology: A contrastive study. In *Proceedings of MT Summit VIII Workshop on Teaching Translation* (pp. 13-17), Santiago de Compostela, Spain.
Kerr, P. (2014). *Translation and own-language activities.* Cambridge University Press.
Klaudy, K. (2003). *Languages in translation: Lectures on the theory, teaching and practice of translation with illustrations in English, French, German, Russian and Hungarian.* Scholastica.
Kliffer, M. D. (2005). An experiment in MT post-editing by a class of intermediate/advanced French majors. In *Proceedings of the 10th EAMT Conference: Practical Applications of Machine Translation* (pp. 160-165), Budapest, Hungary.
Kliffer, M. D. (2008). Post-editing machine translation as an FSL exercise. *Porta Linguarum*, 9, 53-67.
Knospe, Y., Sullivan, K., Malmqvist, A., & Valfridsson, I. (2019). Observing writing and website browsing: Swedish students write L3 German. In E. Lindgren & K. Sullivan (Eds.), *Observing writing: Insights from keystroke logging and handwriting* (pp. 258-284). Brill.
Knowles, C. (2016). *Investigating instructor perceptions of online machine translation and second language acquisition within most commonly taught language courses* [Doctoral dissertation, The University of Memphis]. ProQuest Dissertations Publishing.
Kol, S., Schcolnik, M., & Spector-Cohen, E. (2018). Google Translate in academic writing courses? *The EuroCALL Review 26*(2), 50-57.
La Torre, M. D. (1999). A web-based resource to improve translation skills. *ReCALL*, *11*(3), 41-49.
Larson-Guenette, J. (2013). "It's just reflex now": German language learners' use of online resources. *Die Unterrichtspraxis, 46*(1), 62-74.
Lee, S.-M. (2020). The impact of using machine translation on EFL students' writing. *Computer Assisted Language Learning*, *33*(3), 157-175.
Leow, R. P. (2015). *Explicit learning in the L2 classroom: A student-centered approach*. Routledge.
Leow, R. P., Johnson, E., & Zárate-Sández, G. (2011). Getting a grip on the slippery construct of awareness: Toward a finer-grained methodological perspective. In C. Sanz & R. P. Leow (Eds.), *Implicit and explicit language learning: Conditions, processes, and knowledge in SLA and bilingualism* (pp. 61-72). Georgetown University Press.
Leow, R. P., & Mercer, J. (2014). Depth of processing in L2 learning: Theory, research, and pedagogy. *Journal of Spanish Language Teaching*, *2*(1), 69-82.

Luton, L. (2003). If the computer did my homework, how come I didn't get an "A"? *The French Review, 76*(4), 766-770.

McCarthy, B. (2004). Does online machine translation spell the end of take-home translation assignments? *CALL-EJ Online*, *6*(1).

McLaughlin, B. (1990). Restructuring. *Applied Linguistics, 11*(2), 113-128.

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.

Musk, N. (2014). Avoiding the target language with the help of Google: Managing language choices in gathering information for EFL project work. *TESOL Quarterly*, *48*(1), 110-135.

Németh, G. (2019, October 17). *Machine translation: A short overview*. Towards Data Science, a Medium publication.

Niño, A. (2004, January 6-7). *Recycling MT: A course on foreign language writing via MT post-editing.* [Paper presentation]. CLUK (Computational Linguistics United Kingdom) 7th Annual Research Colloquium, Birmingham, United Kingdom.

Niño, A. (2008). Evaluating the use of machine translation post-editing in the foreign language class. *Computer Assisted Language Learning*, *21*(1), 29-49.

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

Niño, A. (2020). Exploring the use of online machine translation for independent language learning. *Research in Learning and Technology*, 28.

O'Neill, E. M. (2013). Online translator usage in foreign language writing. *Dimension*, 74-88.

O'Neill, E. M. (2014). Real-life technology and the L2 French classroom: Online translation usage among intermediate French students. *Selected Proceedings of the AATF Convention*, 36-42.

O'Neill, E. M. (2016). Measuring the impact of online translation on FL writing scores. *The IALLT Journal of Language Learning Technologies*, *46*(2), 1-39.

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

O'Neill, E. M. (2019b). Training students to use online translators and dictionaries: The impact on second language writing scores. *International Journal of Research Studies in Language Learning*, *8*(2), 47-65.

Osswald, I. (2010). *Examining principled L1 use in the foreign language classroom* [Unpublished master's thesis]. Florida Atlantic University.

Pena, M. I. C. (2011). The potential of digital tools in the language classroom. *International Journal of the Humanities*, *8*(11), 57-68.

Pym, A. (2013). Translation skill-sets in a machine-translation age. *Meta*, *58*(3), 487-503.

Ramsden, T. (2018) Translation in language teaching (TILT): Implementing translation techniques as effective communicative tools in the language learning/teaching environment. *Acta Humanistica et Scientifica Universitatis Sangio Kyotiensis, Humanities Series No. 51*, 249-274.

Richmond, I. M. (1994). Doing it backwards: Using translation software to teach target-language grammaticality. *Computer Assisted Language Learning*, *7*(1), 65-78.

Robinson, P. (1995). Attention, memory, and the "noticing" hypothesis. *Language Learning, 45*(2), 283-331.

Schmidt, R. (1990). The role of consciousness in second language learning. *Applied Linguistics, 11*(2), 129–158.

Selcuk, H., Jones, J., & Vonkova, H. (2019). The use of Google Translate as an ICT tool in web-based collaborative writing: Self-reported accounts of EFL learners. In H. Kratochvilova & R. Kratochvil (Eds.), *Proceedings of IAC 2019 in Vienna* (pp. 74-80). Czech Institute of Academic Education.

Shei, C.-C. (2002a). *Combining translation into the second language and second language learning: An integrated computational approach* [Unpublished doctoral thesis]. University of Edinburgh.

Shei, C.-C. (2002b). Teaching MT through pre-editing: Three case studies. In *Proceedings of the 6th EAMT Workshop on Teaching Machine Translation* (pp. 89-98), Manchester, United Kingdom.

Somers, H. (2001). Three perspectives on MT in the classroom. In *MT Summit VIII Workshop on Teaching Machine Translation* (pp. 25-29), Santiago de Compostela, Spain.

Somers, H. (2003). Machine translation in the classroom. In H. Somers (Ed.), *Computers and Translation. A translator's guide, Vol. 35* (pp. 319–340). John Benjamins.

Somers, H., Gaspari, F., & Niño, A. (2006). Detecting inappropriate use of free online machine-translation by language students: A special case of plagiarism detection. In *Proceedings of the 11th Annual Conference of the European Association for Machine Translation* (pp. 41-48), Oslo, Norway.

Stapleton, P. (2005). Using the web as a research source: Implications for L2 academic writing. *The Modern Language Journal*, *89*(2), 177-189.

Stapleton, P., & Ka Kin, B. L. (2019). Assessing the accuracy and teacher's impressions of Google Translate: A study of primary L2 writers in Hong Kong. *English For Specific Purposes*, 56, 18-34.

Steding, S. (2009). Machine translation in the German classroom: Detection, reaction, prevention. *Die Unterrichtspraxis 42*(2), 178-189.

Sukkhwan, A. (2014). *Students' attitudes and behaviors towards the use of Google Translate* [Unpublished master's thesis]. Prince of Songkla University.

Tomlin, R. S., & Villa, V. (1994). Attention in cognitive science and second language acquisition. *Studies in Second Language Acquisition 16*(2), 183-203.

Tsai, S.-C. (2019). Using Google Translate in EFL drafts: A preliminary investigation. *Computer Assisted Language Learning*, *32*(5-6), 510-526.

Valijärvi, R.-L., & Tarsoly, E. (2019). Language students as critical users of Google Translate: Pitfalls and possibilities. *Practitioner Research in Higher Education*, *12*(1), 61-74.

VanPatten, B. (2004). Input processing in second language acquisition. In B. VanPatten (Ed.), *Processing instruction: Theory, research, and commentary* (pp. 5-31). Lawrence Erlbaum.

Vermes, A. (2010). Translation in foreign language teaching: A brief overview of pros and cons. *Eger Journal of English Studies X*, 83-93.

Vold, E. T. (2018). Using machine-translated texts to generate L3 learners' metalinguistic talk. In A. Haukås, C. Bjørke, & M. Dypedahl, *Metacognition in language learning and teaching* (pp. 67-97). Routledge.

White, K., & Heidrich, E. (2013). Our policies, their text: German language students' strategies with and beliefs about web-based machine translation. *Die Unterrichtspraxis, 46*(2), 230-250.

Williams, L. (2006). Web-based machine translation as a tool for promoting electronic literacy and language awareness. *Foreign Language Annals*, *39*(4), 565-578.

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, L., Gouws, S., Kato, Y., Kudo, T., Kazawa, H., Stevens, K., Kurian, G., Patil, N., Wang, W., Young, C., Smith, J., Riesa, J., Rudnick, A., Vinyals, O., Corrado, G., Hughes, M., & Dean, J. (2016). Google's neural machine translation system: Bridging the gap between human and machine translation. *ArXiv, abs/1609.08144*.

Wuttikrikunlaya, P., Singhasiri, W., & Keyuravong, S. (2018). The use of online tools in L2 writing: A study of Thai university students. *Journal of English Language Teaching and English Linguistics*, *30*(1),107-148.

Xu, M., & Wang, C. (2011). Translation students' use and evaluation of online resources for Chinese-English translation at the word level. *Translation & Interpreting Studies*, *6*(1), 62-86.

Yang, Y., & Wang, X. (2019). Modeling the intention to use machine translation for student translators: An extension of Technology Acceptance Model. *Computers & Education*, 133, 116-126.