Web-Based Machine Translation as a Tool for Promoting Electronic Literacy and Language Awareness

Lawrence Williams
University of North Texas

Abstract: *This article addresses a pervasive problem of concern to teachers of many foreign languages: the use of Web-Based Machine Translation (WBMT) by students who do not understand the complexities of this relatively new tool. Although networked technologies have greatly increased access to many language and communication tools, WBMT is still ineffective for translating text into another language, especially when the user of the software is not able to make grammaticality and acceptability judgments in the target language. This article explains some specific limitations of WBMT (with examples from French) and provides a pedagogical plan for teachers to present this tool to students in order to promote language awareness and electronic literacy, which could help reduce the widespread misuse of this tool by students.*

Key words: *electronic literacy, language awareness, technology-enhanced language learning, tools for language learning, Web-Based Machine Translation*

Language: *French*

Introduction

Since the first half of the 20th century, there has been notable progress in the area of machine translation (MT), particularly since the development of computer hardware and software that are capable of using sophisticated algorithms to produce translated texts of better quality (Bennett, 2000; Carl & Way, 2003; Germann, Jahr, Knight, Marcu, & Yamada, 2004; Güvenir & Cicekli, 1998; Shuttleworth, 2003; Wilks, 2003). During the past century, much work has focused on MT in specific fields, such as chemical safety training (Takala, Pesonen, Kulikov, & Jäppinen, 1991), law enforcement (Hansen, Sorensen, & Johnson, 2002), legal discourse (Lerat, 2002; Kit, Webster, Sin, Pan, & Heng, 2004), patent information (Dollerup, 2002), and so forth. Problems related to translating between languages with different alphabets and writing systems have also received much attention (Herath, Hyodo, Kunieda, & Ikeda,1996; Mankai & Mili, 1995; Tou, 2000; Zantout & Guessoum, 2001), as have issues related to translation of various parts of speech and syntactic structures (Cheong, 1987; Kent & Pitt, 1996; Miller, 2000).

There is little research that examines how MT tools can be used in foreign language education. This is understandable since students are expected to learn

Lawrence Williams (PhD, Pennsylvania State University) is Assistant Professor of French at the University of North Texas, Denton, Texas.

how to communicate in a foreign language, thereby rendering the use of Web-Based Machine Translation (WBMT) superfluous, as typing a text and having the software translate it involve neither communicative activity nor language analysis. Nonetheless, anecdotal evidence points to widespread use of WBMT for homework and writing assignments. Rather than looking only at the possible misuses of this relatively new electronic tool, however, we may wish to examine it further for its potentially positive applications in the study of foreign languages.

The professional literature linking MT and foreign language education has explored the need for MT in the curriculum of translation programs (Lewis, 1997; McCarthy, 2004), comprehensibility and acceptability of translated texts (Leffa, 1994; Petrarca, 2002), the relationship between MT and English as a global language (Cribb, 2000), and techniques for detecting plagiarism and work produced by WBMT software (Luton, 2003). This paper aims to explain WBMT to readers who may not be very familiar with it or who would find it useful to have a model for presenting this tool to students.

Interviews and conversations with colleagues who teach many different languages in a wide variety of educational contexts have revealed that the first reaction to addressing issues of academic dishonesty (e.g., submitting nonoriginal or machine-translated work) is to reprimand students and punish them with a failing grade. While these are often appropriate remedies, the issues related to using WBMT for writing assignments and other tasks could be presented to students at the beginning of a course as a preventative measure. Through careful analysis of translation errors related to one or more specific lexical items and syntactic structures, students will gain a clearer understanding of WBMT, the act of translation, and language in general. They will also have opportunities to develop electronic literacy skills if they analyze, at some level,

product placement, corporate disclaimers, and the discourse of marketing.

Why is electronic literacy important for language learners? According to Hall (2001), "How well we prepare learners of additional languages to meet the social, political, and economic challenges of the next several decades will depend in part on our success in integrating technology into the foreign language curriculum" (p. 60). Hall's advice should not be interpreted as categorical approval for the integration of all types of technology into the foreign language curriculum. Instead, students and teachers need to learn how to evaluate tools in order to understand how, if at all, they might benefit foreign language learning and teaching. Elsewhere, Hall (1999) notes that all domains and modes of communication "are likely to involve not only conventional written and oral modalities but, given the influence of technology in our lives today, electronic ones as well" (p. 38). This assumption is supported, in the United States, by a variety of reports prepared by the Pew Internet and American Life Project (2004). The importance of electronic literacy is also recognized by Canada's National Literacy Secretariat (2006):

> Globalization, new technologies, and new forms of organization in business and government have made the world a more complicated place. It means that Canadians will have to be more skilled at reading, writing, counting and computing to compete for jobs in the new economy, and to take their rightful place as informed citizens in a democracy.

In the context of U.S. foreign language education, the National Standards in Foreign Language Education Project (1999) includes technology as a main strand that serves as a cohesive device in its "'Weave' of Curricular Elements" (p. 33). Kern & Warschauer (2000) remind educators, however, that

> the computer, like any other technological tool used in teaching (e.g.,

pencils and paper, blackboards, over-head projectors, tape recorders), does not in and of itself bring about improvements in learning. We must therefore look to particular *practices of use* in particular contexts (p. 2)

if we, as educators, wish to understand if and how new technologies can improve curriculum, instruction (teaching and learning), and assessment. Shetzer and Warschauer (2000) view electronic literacy as knowledge and skills that involve "what has been called information literacy—the ability to find, organize, and make use of information—but electronic literacy is broader in that it also encompasses how to read and write in a new medium" (p. 173). They divide electronic literacy into three overlapping areas: communication (interpersonal mode of communication, both synchronous and asynchronous); construction (presentational mode of communication); and research (interpretive mode of communication). The various types of knowledge required to use tools for each of these areas is subsumed in each area.

As a broader category than information literacy, electronic literacy includes the evaluation of new information processing and communication tools that have been developed for use in computer- and network-based environments. In the specific case of WMBT, it is true that free, online versions of most software produce many inaccurate, unacceptable translations, and students should therefore be told not to use WMBT as the sole means of writing a composition or doing homework (either because the instructor considers it to be cheating or because it is usually a waste of time). Nonetheless, students should learn how to evaluate WMBT (as well as other tools) from different perspectives or subject positions (Selber, 2004) since the development of multiple literacies involves not only functional literacy (e.g., using technology), but also critical literacy (e.g., questioning technology) and rhetorical literacy (e.g., producing or influencing technology).

These three types of literacy "are meant to be suggestive rather than restrictive, and more complementary than in competition with each other" (p. 24). If students learn how to evaluate WBMT from different perspectives, they will understand better when its use could, if ever, be appropriate. In addition, the process of evaluating WBMT will provide students with a model for critiquing and reflecting on other information processing and communication tools.

In the next section of this article, an analysis is provided of the quality of English–French translations of words, phrases, and sentences produced by three different sites: AltaVista-Babel Fish (AV), Google Language Tools (GO), and Free Translation (FT).[1] This is followed by ideas for lesson plans and guidelines for incorporating both language awareness and electronic literacy into one or more sessions with activities involving WBMT.

Analysis of English–French Translations[2]

The cross-linguistic analysis provided in this section offers teachers of French some parts of speech and structures that can be presented to students in a variety of ways, either all at once or individually, as time and scheduling permit. In certain sections, there are suggestions for expanding the analysis since space is not available in this text to explain in detail many related issues. Although the examples of translations are English–French, teachers of other languages will be able to use this analysis as a model because the categories (e.g., prepositions, nouns, etc.) are based on parts of speech.

Before presenting a selected number of examples of software-produced translations, here is a brief overview of some popular online translation sites for readers who may not be familiar with them.

Since the 1990s, global networking has provided Internet users with access to a variety of free information-processing and communication tools, one type of which is WBMT. As search engines

have grown in popularity, so has advertising and product placement on Web pages with high levels of cyber traffic. Although WBMT software is provided as a free service to visitors of sites and portals such as AOL, AltaVista, Free Translation, Google, La Toile du Québec, Lycos, Sherlock, Voila.fr, and others, these same sites and portals include product information and advertising for translation software that is for sale by the companies that offer free WBMT. Although AltaVista-Babel Fish (AV) and Google Language Tools (GO) have product tie-ins with the same software company, Systran, there are a few rather subtle differences in words and sentences translated by these sites. Nonetheless, in more than 90% of the data collected for this study and others, the translations were identical. Free Translation (FT) is a stand-alone corporate site not associated with a search engine or portal. Instead, the site serves as a marketing tool that promotes other MT products as well as professional (human) translation services available through SDL International, which owns FT.

Prepositions

All three WBMT sites provide accurate translations of the English preposition *to* in most structures. They correctly produce *au* [*to/at/in the*–masculine singular] as the contraction of *à* [*to/at/in*] + *le* [*the*–masculine singular], and they regularly use *en* [*to/at/in*] when it is followed by a sociogeopolitical unit[3] with feminine grammatical gender or one that begins with a vowel (or vowel sound). FT, however, is not as consistent as the other two sites. For example, *to Florida → *à Floride* (instead of *en Floride*), *to Argentina → *à Argentine* (instead of *en Argentine*), *to Spain → *à l'Espagne* (instead of *en Espagne*), *to China → *à la Chine* (instead of *en Chine*). In the first two examples, the geographic units are introduced only by the preposition *à*, which is what would be used before the name of a city. The same structure (*à* + country) appears with the countries Morocco and Denmark, which are both masculine in French and

require the structure (*au* + country), which leads to the conclusion that certain regions, countries, and continents have not been properly identified or labeled as such (i.e., noncities) by the software's programmers.

The English preposition *from* is translated very accurately when its phrasal object is a common noun. In addition, the contraction *du* is produced in appropriate syntactic contexts. There is, however, a major flaw with AV and GO when *from* is followed by a proper noun or, specifically, the name of a geographic region or country. For example, *I am from Denmark* becomes *Je suis le Danemark* [I am Denmark]. Likewise, *She is from Canada* is translated as *Elle est le Canada* [She is Canada], and so forth. FT has a different, albeit just as serious, problem. There is apparently only one sociopolitical unit that is correctly translated following the preposition *from*, namely *from* (the province of) *Quebec → du Québec*. As was mentioned above regarding the preposition *to*, FT has the most problems with geographic units since so many of them have not been identified and labeled as such by the software's programmers.

Adjectives

All three WBMT sites have almost no trouble with adjective–noun agreement, as long as the adjective is adjacent to the noun it modifies. Even when an adjacent adjective is in a phrase set off by commas, it agrees with the noun in number and gender, as seen in the following example containing the (feminine singular) adjective *déçue* (which modifies the feminine singular noun *communauté*): *The scientific community, disappointed with the situation, decided to act → La communauté scientifique, déçue (de) la situation, a décidé d'agir.* When an adjacent adjective precedes the noun it modifies, only AV and GO produce a form with adjective–noun agreement. For example, *Disappointed, the scientific community . . . → Déçue, la communauté scientifique . . .* (AV/GO); *Déçu, . . .* (FT). None of these sites successfully produce agreement between nouns and remote (i.e.,

nonadjacent) adjectives. *Disappointed with the situation, the scientific community . . . → *Déçu (de) la situation, la communauté scientifique . . . .*

Adjective placement enjoys the same high rate of success and output quality as adjective–noun agreement. All three sites deal well with most of the French adjectives that change meaning depending on their (pre- or postnoun) placement. One notable exception is the translation of *old*, which can mean both former and aged in English. The software cannot discern the semantic properties of *old* based on context; therefore, *old* is always translated as a form of *vieux* [old/aged], while *former* is correctly recognized as *ancien*. Although these translations are certainly acceptable, the average student might not take into consideration the polysemic nature of the English adjective *old*. This example, like others, highlights the importance of students understanding that computer software often cannot make important meaning-related distinctions.

Nouns

For most common sports mentioned in first- and second-year textbooks, the translation works fine in AV and GO. The preposition *à* is even incorporated correctly into most sentences in which forms of the verb *jouer* [to play] are used (in this context). For example, *I play tennis* is translated as *Je joue au tennis*. Incidentally, FT translates this same sentence as **Je joue le tennis* [I play the tennis], an almost literal yet inaccurate translation since French requires a form of the preposition *à* (in addition to the definite determiner *le, la,* or *les*) after a form of *jouer* when referring to playing a game or sport. This same structure is produced by FT for similar sentences, regardless of the sport in the original sentence. The sport of golf, however, might not be in the lexicon of the AV and GO software (or it might not be designated as a sport) since *I play golf* becomes **Je joue le golf* (instead of . . . *au golf*).

Interestingly, both *football* and *soccer* are translated by AV and GO as *football*, and the aspirate status of *h* at the beginning of *le hockey* is not recognized by the software. *Ping-pong* is translated correctly (*Je joue au ping-pong*) only if it is hyphenated. Both *baseball* and *softball* are translated as *baseball* by all three WBMT sites, even though there exist several important differences between these two bat-and-ball sports.[4] For games and sports that are plural, the partitive *des* is produced by FT (*I play chess → *Je joue des échecs,* instead of *Je joue aux échecs*).

Although the syntax of *jouer* (when used to express playing a game or sport) often translates correctly in AV and GO, this is not the case when *jouer* refers to playing a musical instrument. Instead of incorporating the preposition *de* into such sentences, no preposition follows the verb—only a definite determiner precedes the name of the musical instrument. For example, *I play the guitar* becomes **Je joue la guitare* (instead of *Je joue de la guitare*).

In AV and GO, most of the musical instruments are at least in the software's lexicon, and the correct grammatical gender has, for the most part, been assigned to them. However, *I play the clarinet* is translated as **Je joue le Clarinet* by AV and GO, and as **Je joue la clarinette* (instead of *Je joue de la clarinette*) by FT. In the first case (AV/GO), the musical instrument has been transformed into a masculine proper noun and the spelling is not correct. The FT translation has the correct grammatical gender and spelling of *clarinette*, yet the preposition *de* is nonetheless missing. Likewise, *I play the oboe* is simply reproduced by AV and GO as . . . *l'oboe,* suggesting that the software does not contain this word. FT produces the correct noun, but without granting it word-initial *h* aspirate status (**l'hautbois* instead of *le hautbois*). The *flute* is translated by AV and GO not as *la flûte*, but as *la cannelure,* a technical term that can be the French equivalent of the English word *flute* when used in fields such as architecture, botany, and geology.

When *the harp* is translated, the AV and GO software do not recognize the aspirate status of the *h*, resulting in **l'harpe* instead of *la harpe,* which is produced by FT.

There is a wide discrepancy between AV/GO and FT regarding the aspirate status of the word-initial *h*. This can be shown to students by comparing, for example, *hibou* (correctly identified as *le hibou* [*the owl*] by both AV/GO and FT) and *Hongrie* [the country Hungary] (aspirate *h* in AV/GO, but not in FT). The discovery of different treatments of the aspirate *h* is an excellent starting point for a lesson or series of lessons on the word-initial *h* in French, as well as other phonetic/phonological phenomena. Teachers of French can also introduce the social controversy surrounding the widespread rumor that the word-initial *h* in *haricot vert* [green bean] has been declared or accepted as nonaspirate by the *Académie française.*[5]

Verbs and Verb Phrases

French verbs (like those in many other European languages) are extremely complicated since they express mood, tense, and aspect, and they carry markers of grammatical person and number (and, in certain forms, gender). While AV/GO and FT deal relatively well with both simple and compound forms, tense and aspect do not transfer or translate directly when concepts such as *ago, have/has/had just,* and *for* (when referring to a certain amount of time) are used in sentences. The expression *have/has/had just* includes an additional level of complexity since many North American speakers of English tend to use *just* + verb to express having just done something (the same structure used when *just* means simply or only) instead of *have/has just* + verb. None of the WBMT sites successfully translate *have/has/had just* + verb correctly into a form of *venir* + *de* + infinitive.

In Table 1, it is clear that both AV/GO and FT have difficulties translating the concept *ago*. Both WBMT sites produce an incorrect or unacceptable version of the first sentence (S1), but the errors are

related to different phenomena. In the case of AV/GO, the singular noun neighbor is translated as a (feminine) plural noun, *voisines*. Moreover, the syntax is jumbled: the noun phrase (direct object) *nos voisines* has been separated by *trois heures,* which in turn is separated from the first half of its adverbial complement, *il y a.* However, in the FT translation of S1, the syntax of the adverbial complement itself is simply reversed, which reflects the English syntax of *three hours ago* (**trois heures il y a* instead of *il y a trois heures*).

Table 2 demonstrates how simply modifying one small part of a word (in this case, making a noun plural) can result in a change in the quality of the translated sen-

TABLE 1

Translation of *We saw our neighbor three hours ago.*

English S1	We saw our neighbor three hours ago.
AV/GO S1	*Nous avons vu il y a nos trois heures voisines.
FT S1	*Nous avons vu notre voisin trois heures il y a.
Acceptable S1	Nous [avons vu] notre voisin(e) [il y a] trois heures. We [saw/have seen] our neighbor [ago] three hours.

TABLE 2

Translation of *We saw our neighbors three hours ago.*

English S2	We saw our neighbors three hours ago.
AV/GO S2	Nous avons vu nos voisins il y a trois heures.
FT S2	*Nous avons vu nos voisins trois heures il y a.
Acceptable S2	Nous [avons vu] nos voisin(e)s [il y a] trois heures. We [saw/have seen] our neighbors [ago] three hours.

tence. In the second sentence (S2), which includes the same use of *ago* as the sentence in Table 1, the plural noun *neighbors* was used instead of the singular and AV/GO produced a correct or acceptable version of the original. However, this change resulted in no improvement for FT, which still had difficulty with the syntax of the adverbial complement.

Even with verbs and verb phrases with no problems related to syntax, translation can be problematic. Such is the case of the verb *to know,* which is often translated as *savoir* or *connaître* in French. (Teachers of other Romance languages will certainly recognize these as counterparts to *saber* and *conocer* (Spanish), *sapere* and *conoscere* (Italian), and so forth.) The AV software treats these two distinct verbs correctly at a very general level. It distinguishes between knowing the answer (*savoir*) and knowing, for example, a person, a town, or a restaurant (*connaître*). It also correctly translates knowing how to do something. *I know how to swim.* → *Je sais nager.* However, there are some peculiarities involving accusative case (direct object) pronouns. Consider the following sentences: *I know him. I know her. I know them.* → *Je le connais. Je la connais. *Je les sais.* The introduction of the third-person plural accusative (direct object) pronoun triggers a use of a form of *savoir,* but the random switch between forms of *savoir* and *connaître* does not occur when the direct object is in nonpronominal form in postverbal position.

Another type of problem associated with verbs and verb phrases is that of pronominal verbs, which have what amounts to in many cases an unanalyzable object pronoun that matches the grammatical person and number of the subject. These are often presented in first- and second-year French textbooks when students are learning or reviewing the grammar and vocabulary of the daily routine.

The AV/GO software performs moderately well with the reflexive pronominal verbs *se laver* [to wash], *se réveiller* [to wake up], and *se brosser* [to brush].

However, forms of *to fall asleep* are translated as forms of *tomber endormi* [*to fall asleep*—a literal, yet inaccurate translation] instead of *s'endormir,* which is used correctly by the FT software. Nonetheless, FT has its own set of problems, which are primarily related to its difficulty dealing with particle verbs (e.g., to turn in, to turn down, to turn out, to fill in, to fill out, to fill up, etc.), a common shortcoming of most free WBMT sites. For a verb such as *to wake up,* the particle *up* can be adjacent or remote. When *up* is separated from the verb by a direct object, the software treats up as an isolated preposition, producing the translation *en haut* (which could mean *upstairs, up high, on high,* etc.). FT translates *She woke up* and *She woke up the children* correctly (as *Elle s'est réveillée* and *Elle a réveillé les enfants*), yet *She woke the children up* (with the particle separated from the verb by a direct object) becomes **Elle est réveillée les enfants en haut* [*She is woken the children upstairs]. The AV/GO software, however, correctly translates forms of *to wake up,* regardless of the location of the particle.

After reading a few essays produced by WBMT sites, any teacher will begin to recognize that AV and GO use *vers le haut* [toward the top, upward] when *up* is treated as an isolated preposition, just as FT treats *up* as *en haut.* Likewise, AV and GO produce *vers le bas* [toward the bottom, downward] when *down* is considered an isolated element instead of a verbal particle. FT, on the other hand, can analyze *down* correctly at some level, producing **J'ai refusé la chaleur* [I refused the heat] as the translation of *I turned down the heat.* Since the software does not make all lexical choices based on the context of the sentence, it recognizes as the default verb a form of *refuser* [to refuse/to turn down] instead of a form of *baisser* [to lower/to turn down]. The problems of semantic conflict and lexical choice are also highlighted by the use of *la chaleur* [heat/temperature] instead of *le chauffage* [heat/heating/heating system].

An exercise based on the translation of particle verbs can be one of the most

powerful and effective ways to demonstrate to foreign language learners cross-linguistic and, by extension, cross-cultural differences, especially since verb forms and syntax are so different when comparing English and French. Pronominal (reflexive and reciprocal) French verbs play an important role in the communication of everyday routines and events, and many of these are particle verbs in English. If students can see that communicating in another language is not simply a matter of plugging words into a formula that can be calculated by a machine, they will begin to understand language and communication as complex and multilayered.

Pedagogical Plan for Presenting and Explaining WBMT to Students

Search engines and portals can be presented as prime locations for MT product placement due to the high volume of cyber traffic. Before directing students to one or more search engine sites or portals, teachers should ask students to provide names of popular search engines. The students will most likely mention Google, Yahoo!, AOL, and others. In order to incorporate opportunities to develop electronic literacy awareness, the whole class can be asked to determine a working definition of a search engine and/or portal. Once the students have collectively agreed on the main features and elements of their definition, the teacher should open an online dictionary or specialized site containing explanations of terms related to newer technologies, such as Webopedia or TechEncyclopedia.[6] Next, the teacher (or one of the students) can open a site such as Google in order to access the free WBMT service, which can be found in the Language Tools section.

Time permitting, the students can first analyze the different zones on the welcome page of one or more portals or search engines. A portal (e.g., Voila.fr, Yahoo!) will feature links to the weather and shopping, as well as thematically organized links, chat rooms, free e-mail, and at least a few advertisements. A site that serves primarily as a search engine (e.g., Google) will not have as many extra features. The Voila.fr portal[7] has main feature categories grouped together: communication (*Communiquer*), advertising, news (*L'essentiel aujourd'hui*), and searching; within each of these zones there are various products and services. For example, *t'Chat* [chat], *E-mail,* and *Traducteur* [translator] are in *Communiquer* [communicating], while *Météo* [weather], *Dernière minute* [latest news], and *Le journal* [newspaper/news presentation] are in *L'essentiel aujourd'hui* [what you need (to know) today]. Although other portals may not have products and services grouped in the same way, students can decide how the organization of the page is intended to help visitors navigate the site. If they do not notice user-friendly elements of the design, they can be guided by the teacher to observe the use of different font sizes, colors, and physical space between zones.

Once the students have found the link *Traducteur*, they can go to the WBMT site. There are two very important links in the section *Plus d'infos* [more information], and the pages found there do indeed contain much more information about this language tool. The *Conditions d'utilisation* [conditions of use] page explains, among other things, that the quality of MT cannot be guaranteed and that Voila.fr/Systran cannot ensure customer satisfaction for this free service. The *Aide* [help] page (within *Plus d'infos*) explains how to translate a Web page versus text. There is also a reminder that spelling and punctuation are important, since the software can deal with source information only as well as it is entered by the user. After navigating through the portal and WBMT site, students should have an opportunity to discuss product placement, both the potential benefit to companies involved and how consumers (people who visit this site) might be affected. Other questions should deal with the disclaimers that the students have seen. These can be posed either while they are reading them or after the class has

had a chance to visit the entire *Traducteur* site.

If there is little or no class time to dedicate to this type of lesson using a whole-class participation structure, the following questions can be given to students as an assignment to be done at home, which could be followed by a brief lesson in class when students would present their findings. Likewise, if the topic of WBMT were discussed briefly in class before assigning some or all of the questions to students, brief written reports could be submitted by each student or groups of students as a way to summarize their findings.

Item 1
Visit an online search engine or portal site such as Google, Yahoo!, or Voila.fr and explain where on the page the link to the translation tool is located. Is the link easy to find, or is it not even on the main page of the search engine or portal site? (Students might already know that links are strategically placed on Web pages, but a concrete example might help illustrate the importance of link and product placement. Several years ago, the link to the free translation tool on AltaVista's site used to be at the very bottom of the welcome page; however, as this service has become increasingly popular, the link to *Babel Fish Translation* has gradually moved up. It is now prominently displayed in its own row directly under the main search box.)

Item 2
What other types of free information-processing or communication services does this site provide?

Item 3
Is there an advertisement for a product made by the company that provides the free online translation service? Does the advertisement explain why people might be interested in purchasing a product as opposed to using the online service?

Item 4
Can you find any disclaimers regarding the quality of the translations produced by the free online software? Are they available on the main page of the translation area of the site? What kinds of explanations are given in disclaimers about free online translation software? (If students compare WBMT sites, they will realize that sometimes a link to a disclaimer is not provided on the main page of the site, and other times it can only be found after at least one word or phrase has been translated, which is the case for GO. On the AV site, an obvious disclaimer is nowhere to be found, but some caveats are made in the *Help* section.)

Item 5
In addition to typing text into a box and having the software translate it, it is also possible to have the online software translate a Web site. Find an English-language site and have the software translate it into French, then do the same thing with a French-language site and translate it into English. Take some time to read through a part of each translated site. Are there any errors in the site translated into English? Even if there are errors, can you figure out the main idea of the site translated into English? What about the site translated into French? Are you able to spot errors as easily? (The purpose of this series of questions is to reinforce the idea that the free online software is intended to be used by people who can make judgments in the language of the translated text or by people who take the time to look up and check words and expressions if they are not able to make such judgments.)

There are certainly many different ways to approach this electronic literacy exercise. It is most important to remember here that the focus is on guiding students through different Web sites and helping them become more aware of how content is organized, labeled, and presented in certain areas of cyberspace. This type of analysis of WBMT will also serve as a model for evalu-

ating online resources and other tools used by electronically literate students.

In addition to learning about the tool's placement on Web sites and the marketing of its services, students should be familiar with the many linguistic limitations of WBMT. This second part of the lesson involves raising students' awareness of basic problems associated with translation and language use. For example, although a text produced by WBMT software may be incomprehensible, it is possible to determine, in many cases, the reasons for the mistranslation(s). Although the terms *polysemy, lexical ambiguity,* and *structural ambiguity* might not mean much to students, examples in English can quite easily illustrate the points that polysemy can create lexical ambiguity, and word order can determine the level of structural ambiguity in a sentence.[8] This part of the explanation should force students to think about language as a communication tool, not as a set of decontextualized vocabulary words or phrases. The English word *speaker* is a good example to use when considering polysemy since it has at least two common meanings: a person who is talking or an apparatus that transmits sound. In order to see if students understand this concept, have them create a list of similar words (not limited to nouns) either in class or at home as preparation for or a continuation of this lesson on WBMT software.

Once students have thought about lexical and structural ambiguity, they might be able to make some predictions about the potential problems involved in using WBMT. Using one of the WBMT sites, it would be helpful to start with single words to show students how the software determines which word is its default form. For example, AV translates the English word *tax* as *impôt*. Although this French word is associated with a tax on income, *taxe* would be used more often to refer to a tax on the sale of goods and services. Another phenomenon worth exploring with students is that a misspelled word (in the original text) is not translated—it is simply reproduced.

Depending on the amount of time available, the teacher can ask students to think of other examples, in class or as homework, that might illustrate this challenge faced by human translators and WBMT software. Students should be asked to guess which word would be the default in French, explain why, then compare their guesses to what they find using one of the WBMT sites. At some point during this explanation of WBMT, students should have a chance to analyze sentences that have been translated both accurately and inaccurately. The MT literature suggests what many language teachers certainly already know: verbs and prepositions seem to be the most problematic elements in texts produced by MT software. If there is not enough time to focus on both verbs (especially pronominal verbs) and prepositions, one or the other should be presented and explained to students.

If little or no time is available for a class lesson on the linguistic limitations of WBMT, some or all of the questions provided below can be given to students as an at-home assignment.

Item 6

What is the software's translation of *car?* Use one or more dictionaries to see if there are any other words that mean *car* since the online software provides only one solution. You could also try this with English words that can be a noun or a verb, such as *drive.* See how the software translates *drive* or a similar word in isolation versus in a sentence. Is there a difference? Explain the problem. (This series of questions will force the students to ask themselves why one word was chosen and not another in the case of words that do indeed have more than one possible translation. The goal of this type of reflection on language is to dispel a somewhat commonly held belief that words and concepts have a one-to-one correspondence between or among languages.)

Item 7

How does the software deal with gender for nouns that refer to people, such as *neighbor?* Other examples would be *cousin, mother, father, sister,* etc. Is the software able to identify the correct translated form for some but not others? One way to test this is by using some of these words in isolation and then using them in sentences; by submitting sentences such as *My neighbor is intelligent* and *My neighbor is an intelligent girl,* students can determine if the translation of *neighbor*—the masculine or feminine form—depends on the context of the sentence. (This series of questions is relevant for languages in which nouns that refer to people have feminine forms. A related task could involve adjective–noun agreement involving female first names. Students who try a variety of female first names in sentences with adjectives will realize that only a certain number of traditional female first names are recognized by the software as feminine. This can reinforce the fact that accurate translations in such structures can occur only when nouns—both common and proper—have been entered into the software's lexicon and labeled correctly.)

Item 8

What types of problems would you expect with English verbs that have a preposition immediately following them, such as *turn up, turn down, turn in,* etc.? When the preposition—*up, down, in*—is added to the verb *turn*, it changes the meaning entirely. These types of verbs are problematic because the two words must be translated together and even the same verb can have different meanings. Think of a few sentences using *turn down* with different meanings (*lower* and *refuse,* for example), then translate these sentences using the online software. What kinds of problems do you notice with the translations of these verbs?

Other ideas for this lesson plan can be generated from issues that students raise in class or from the analysis of English–French translations produced by AV, GO, and FT that was provided earlier in this article.

Exercises for students in their third or fourth years (or beyond) of language study could be similar, but such students might also be interested in the choices made by the software companies when labeling and coding the entries in the lexicon. A study of female first names and their agreement with adjectives, for example, could have as its point of departure an analysis of the types of English female first names that have been labeled as feminine by the software's programmers. After establishing a list of first names of this type that trigger adjective agreement, the students could then take the study in a different direction by comparing popular female first names in countries where French (or German, Portuguese, etc.) is spoken. Students in a translation class could write to different software manufacturers to inquire about the decision-making process in order to understand how companies prioritize their work, such as labeling nouns as feminine or masculine (or other types of coding and labeling) in the software's lexicon. Regardless of the type of activity students are asked to do, this type of task should be presented as a model for evaluating a tool in order to understand how it has been produced and marketed and how it works.

Conclusion

After viewing and using WBMT tools, students should become aware that the intended user of WBMT software is not the foreign language student. Instead, WBMT is intended, or is at least most effectively used, to help people understand the basic information found on a Web page published in another language. In other words, anglophone students need to realize that they would have a much easier time making sense of a foreign-language Web page translated into "bad" English than an English text translated into bad French. Their ability to judge the quality of the translated French is limited, and understandably so. Students should also be encouraged to consider the importance of words and expressions that are commonly used versus those

that are rare. This could encourage them to think about the frequency of lexical items and structures in their own first and second language use. As long as students are encouraged and guided toward active engagement in some type of reflection on language and language use, they will have opportunities to question and perhaps reevaluate their beliefs and practices regarding communication and new tools at their disposal. In the case of WBMT, students should at least begin to understand in what ways and to what extent this tool may not be suitable for their homework and writing assignments.

Teaching students how to analyze and understand one language tool, such as WBMT software, can serve as a model for learning (or relearning) how to use other tools, such as dictionaries (print-based or electronic; monolingual or bilingual), word-processing software, databases, and so forth. WBMT software is far from being an ideal language tool for students. Nonetheless, students should understand how this tool works and its possible strengths and limitations in different linguistic, social, and educational contexts. Kress (2003) warns that "it is no longer possible to think about literacy in isolation from a vast array of social, technological and economic factors" (p. 1). Shetzer & Warschauer (2000) see literacy as "a shifting target" that educators need to follow closely since "we have to prepare students for their future rather than for our past" (p. 172). If we truly want to prepare foreign language students for life in the 21st century, it is necessary to teach them not only where to find new tools but also how to use—and avoid misusing—them.

Notes

1. AltaVista-Babel Fish can be found at http://babelfish.altavista.com; Google Language Tools can be found at http://www.google.com/language_tools; Free Translation can be found at http://www.freetranslation.com.

2. An asterisk (*) is placed at the beginning of a sentence or group of words in which one or more element is considered ungrammatical and/or unacceptable.

3. Within this analysis of prepositions and geography, it should be noted that U.S. translates correctly as *États-Unis* (with the optional acute accent on the first letter of *États*) in AV and GO, only if there is no space between the first period and the letter S in the abbreviation. When there is a space between the first period and the letter S, this grouping is not translated as a unit. It is simply kept as typed in English and preceded by the preposition *à*. It is also translated correctly (without the optional acute accent on the first letter of *Etats*) if *United States* is used. Incidentally, translated text in FT never includes the optional (on upper-case letters) acute accent.

4. A discussion of English sports terminology used and adopted by the francophone world is beyond the scope of this article. However, the reader may be interested to know that the French term *soft-ball* is widely used by francophones (who play or make reference to this sport) in North America and Europe. For a variety of reasons, there is some confusion and divergence regarding its grammatical gender. Although foreign words (and similar names of sports) that enter French often default to masculine, one French translation of the English word *ball* is *balle*, which is feminine and may be one reason for some of the divergence.

5. The rather contemptuous statement by the *Académie française* on this matter can be found at the following URL: http://www.academie-francaise.fr/langue/questions.html#haricot. A related topic demonstrating phonetic/phonological variation is the *h* in *le hockey*, which is consistently treated as aspirate in European French, yet this is not always the case in Canadian French.

6. Webopedia is located at http://www.webopedia.com/. TechEncyclopedia can be found at http://www.techweb.com/encyclopedia/.

7. Voila.fr is a French-language portal site and has a product tie-in agreement with Systran. The URL for Voila.fr is http://www.voila.fr.

8. Lexical ambiguity can be illustrated by the difficulties in determining the meaning of *pen* in the following sentence: "They are trying to design a better pen. ('writing implement' or 'animal enclosure'?)" (Somers, 2003, p. 124). Structural ambiguity arises when trying to determine, in this sentence, the syntactic function of *yesterday*: "The minister stated that the proposal was rejected yesterday" (p. 125).

References

Bennett, W. S. (2000). Taking the babble out of Babel Fish. *Language International, 12*(3), 20–21.

Carl, M., & Way, A. (Eds.). (2003). *Recent advances in example-based machine translation.* Dordrecht, Netherlands: Kluwer.

Cavalier, Tim. (2001). Perspectives on machine translation of patent information. *World Patent Information, 23,* 367–371.

Cheong, T. L. (1987). The computer translation of interrogatives from English to Malay. *RELC Journal, 18,* 1–18.

Cribb, V. M. (2000). Machine translation: The alternative for the 21st century? *TESOL Quarterly, 34,* 560–569.

Dollerup, Cay. (2002). Translation in the global age. *Language International, 14,* 40–43.

Farghaly, A., & Senellart, J. (2003). *Intuitive Coding of the Arabic Lexicon.* MT Summit IX. Retrieved June 2, 2004, from http://www.systransoft.com/Technology/2003_MTIX_AR.pdf.

Germann, U., Jahr, M., Knight, K., Marcu, D., & Yamada, K. (2004). Fast and optimal decoding for machine translation. *Artificial Intelligence, 154,* 127–143.

Güvenir, H. A., & Cicekli, I. (1998). Learning translation templates from examples. *Information Systems, 23,* 353–363.

Hall, J. K. (1999). The communication standards. In J. K. Phillips (Ed.), *Foreign language standards: Linking research, theories, and practices* (pp. 15–56). Lincolnwood, IL: National Textbook Company.

Hall, J. K. (2001). *Methods of teaching foreign languages: Creating a community of learners in the classroom.* Upper Saddle River, NJ: Prentice Hall.

Hansen, I. G., Sorensen, H. S., & Johnson, E. (2002). Multilingual police communication. *LSP and Professional Communication, 2*(1), 84–97.

Herath, A., Hyodo, Y., Kunieda, Y., & Ikeda, T. (1996). Bunsetsu-based Japanese-Sinhalese translation system. *Information Sciences, 90,* 303–319.

Kent, S., & Pitt, J. (1996). Feature-based & model-based semantics for English, French and German verb phrases. *Language Sciences, 18,* 339–362.

Kern, R., & Warschauer, M. (2000). Introduction: Theory and practice of network-based language teaching. In M. Warschauer & R. Kern (Eds.), *Network-based language teaching: Concepts and practice* (pp. 1–19). Cambridge: Cambridge University Press.

Kit, C., Webster, J., Sin, K. K., Pan, H., & Heng, L. (2004). Clause alignment for Hong Kong legal texts: A lexical-based approach. *International Journal of Corpus Linguistics, 9,* 29–51.

Kress, Gunther. (2003). *Literacy in the new media age.* London: Routledge.

Leffa, V. J. (1994). Machine-translated text: Is it comprehensible to proficient readers? *System, 22,* 391–399.

Lerat, Pierre. (2002). Legal vocabulary and the schemata of legal arguments. *Meta, 47,* 155–162.

Lewis, D. (1997). Machine translation in a modern languages curriculum. *Computer Assisted Language Learning, 10,* 255–271.

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

Mankai, C., & Mili, A. (1995). Machine translation from Arabic to English and French. *Information Sciences—Applications, 3,* 91–109.

McCarthy, B. (2004). Does online machine translation spell the end of take-home translation assignments? [online]. *CALL-EJ Online, 6*(1). Available: http://www.tell.is.ritsumei.ac.jp/callejonline/journal/6-1/mccarthy.html.

Miller, K. (2000). The lexical choice of prepositions in machine translation. *Dissertation Abstracts International, 61* (07), 2685A. (UMI No. 9978132).

National Literacy Secretariat (Canada). Retrieved November 8, 2006 from http://www.hrsdc.gc.ca/en/hip/lld/nls/About/aboutus.shtml.

National Standards in Foreign Language Education Project. (1999). *Standards for foreign language learning in the 21st century.* Lawrence, KS: Allen Press.

Petrarca, M. (2002). Machine translation: A tool for understanding linguistic challenges facing the second language student. *Dissertation Abstracts International, 63* (01), 120A. (UMI No. 3037900).

Pew Internet & American Life Project. Retrieved September 1, 2004 from http://www.pewinternet.org.

Selber, S. A. (2004). *Multiliteracies for a digital age.* Carbondale, IL: Southern Illinois University Press.

Shetzer, H., & Warschauer, M. (2000). An electronic literacy approach to network-based language teaching. In M. Warschauer & R. Kern (Eds.), *Network-based language teaching: Concepts and practice* (pp. 171–185). Cambridge: Cambridge University Press.

Shuttleworth, M. (2003). Translation technology: Myths and reality. *Cahiers AFLS, 9,* 7–18.

Somers, H. (Ed.). (2003). *Computers and translation: A translator's guide.* Amsterdam: John Benjamins.

Takala, J., Pesonen, J., Kulikov, L., & Jäppinen, H. (1991). Machine translation for chemical safety inspection. *Journal of Hazardous Materials, 27,* 213–230.

Tou, J. T. (2000). An intelligent full-text Chinese–English translation system. *Information Sciences, 125,* 1–18.

Warschauer, M., & Kern, R. (Eds.) (2000). *Network-based language teaching: Concepts and practice.* Cambridge: Cambridge University Press.

Wilks, Y. A. (2003). *Machine translation: Its scope and limits.* Cambridge: Cambridge University Press.

Zantout, R. N., & Guessoum, A. A. (2001). An automatic English–Arabic HTML page translation system. *Journal of Network and Computer Applications, 24,* 333–357.