A STUDY OF FACTORS INFLUENCING MACHINE TRANSLATION COMPREHENSION

Mahesh Vanjani, Texas Southern University
vanjanim@tsu.edu

Milam Aiken, University of Mississippi
maiken@bus.olemiss.edu

Mina Park, Southern Connecticut State University
Parkm2@southernct.edu

ABSTRACT

Keywords: Machine Translation, Foreign Language Translation, Translation Comprehension

INTRODUCTION

Every day, millions of text passages are translated from one language to another by machine because of these systems' high speed and low cost compared to human interpretation (Cajvaneau, 2002). This machine translation often results in less than perfect accuracy and concomitant poor human comprehension. As a result, one study (Harzing, Koster & Magner, 2011) found that 50% of a survey's respondents thought the technology was useless beyond getting a very basic understanding of the topic of the translated materials.

Several studies have shown varying degrees of comprehension of translated text among readers. For example, Table 1 summarizes the results of experiments involving multilingual groups exchanging comments with automatic translation provided by the meeting software, and reported understanding varied from 40% to 100%. Text quality varied among the studies, and some translations were undoubtedly better than others. For example, some evidence suggests that western European languages translate to English well because of their similarities (Aiken & Balan, 2011). But, there can be human differences in addition to lingual, text, and software dissimilarities. If we can ascertain what factors affect understanding of mistranslations, we might better determine which potential users of translation software could benefit most.

Table 1. Varying Amounts of Comprehension of Translated Text
(Source: Aiken, 2008)

Study	Group Members	Understood Accuracy
1	24 English in 4 groups. Facilitator added Spanish comments	Spanish->English: 100%
2	3 Spanish, 6 English	English->Spanish: 96% Spanish->English: 85%
3	3 Spanish, 5 English	NA
4	2 Spanish, 2 English	English->Spanish:91%

		Spanish->English: 81%
5	1 German, 1 French, 2 English	German->English: 100%
		German->French: 40%
		French->German: 40%
		French->English: 60%
		English->German: 80%
		English->French: 80%
6	2 English, 2 Chinese	English->Chinese: 93%
		Chinese->English: 100%
7	Translation of 2 English-only- meeting comments. 3 Spanish evaluators	English->Spanish: 72%-74%
8	Translation of 2 English-only- meeting comments. 3 Chinese, 3 Japanese, 3 Korean evaluators	English->Chinese: 43%-74%
		English ->Japanese: 75%-78%
		English->Korean: 55%-78%

Humans have widely varying skills and abilities, and they comprehend text differently, especially when the passage contains numerous word choice or spelling errors (Schiefele, 1999). But, it is not clear exactly what factors influence this understanding. For example, it is reasonable to assume that people who know their native language extremely well might be more able to obtain the meaning of text despite numerous grammatical and word-choice errors. Also, those who are more knowledgeable about the subject matter might be able to extract recognizable key nouns and phrases, leading to a better grasp of the material. Experience with a different language might help (Gottardo & Mueller, 2009). For example, if a person knows some Spanish, even if the translation is in English, he or she might be able to understand garbled text because of previous exposure to translation errors. Finally, those with greater motivation to understand might try harder and spend more time trying to detect meaning.

In this paper, we explore some of these variables with a survey of English speakers exposed to poorly translated text. As in previous studies, results are somewhat inconclusive, necessitating further research on this elusive concept.

MACHINE TRANSLATION ERRORS

People can miscomprehend a passage of text even if the translation is perfect, for example, when the original text is unclear or has errors. With the addition of translation errors, however, comprehension can become even more difficult. Several types of errors can occur with machine translation, and Koponen (2010) describes five possible classifications:

1. **Omitted concept.** A word or words are missing in the translation that consequently leave out some meaning.
2. **Added concept.** A word or words are added in the translation that bring extra meaning not present in the source text.
3. **Untranslated concept.** A word or words from the original text appear verbatim in the target text with no change in language.
4. **Mistranslated concept**. A word or words are translated to the target language but convey the wrong meaning.

5.   **Substituted concept.** The translation is not identical, but can be considered a valid replacement for the original meaning.

A more detailed list of possible machine translation errors includes (Irvine, Morgan, Carpuat, Daumé & Munteanu, 2013; Popović & Ney, 2011; Vilar, Xu, d'Haro & Ney, 2006):

1. **Transliterated words**. If a word is not recognized because it is misspelled, slang, an acronym, outside the dictionary, or for another reason, a word can be injected literally into the translation. For example, "My nme is Adam" could be translated to "Meine nme ist Adam" in German because of the incorrect spelling of "name."

2. **Wrong meaning of word chosen**. Many languages have words with different meanings, and typically, the correct choice is determined by the context of the surrounding words. For example, a key used on a piano is "tecla" in Spanish, but a key used for a lock is "llave." Also, a proper noun could be translated literally, e.g. "Mr. Street" might be translated to "Herr Straße" in German instead of "Herr Street."

3. **Word order rearranged**. If a sentence is complex with several clauses, the word order might not be correct. Translation among western Europeans languages with similar grammatical structures is often good, but machine translations between radically different languages (e.g., English to Chinese) can result in more errors (Linh, 2013).

4. **Other grammatical errors**. Other grammatical errors can include changes in verb tense, lack of parallel structure, lack of subject/verb agreement, vague pronoun references, incorrect punctuation, etc.

5. **Idioms**. Idioms are often used in informal discussions, but are difficult to translate. For example, the equivalent for "pulling my leg" as in kidding or joking is "tomar el pelo" in Spanish, which means literally "to take the hair." However, many translation systems are not sophisticated enough to recognize equivalent idioms and instead translate each word literally, possibly causing confusion.

6. **Poor phrasing**. Some translations might be technically correct, but could be worded better. For example, a translation of "Ich heiße John" from German literally means "I am called John," but better phrasing to English would be "My name is John."

7. **Miscellaneous errors**. Several other errors can occur including use of informal addresses instead of formal (e.g., using "du" instead of "Sie") in German. While the meaning might be understood, the relationships among the principle actors in the passage might not be clear.

The following provides more examples of possible machine translation errors using Google Translate (http://translate.google.com):

1. The Universal Declaration of Human Rights in Chinese (http://www.omniglot.com) was translated to "Human beings are born free in dignity and rights are equal. They are endowed with reason and conscience and should act in a spirit of brotherhood treat each other" but the correct translation should be "All human beings are born free and equal in dignity and rights. They are endowed with reason and conscience and should act towards one another in a spirit of brotherhood." Thus, most of the meaning is conveyed, but there are a few words in the wrong order, causing confusion.

2. In this machine translation from Japanese (https://www.yaqs.co.jp/help/sample5): "If something unforeseen circumstances occur, please immediately inform the Company's employees," the word "something" is added unnecessarily.

3. This machine translation from Japanese (https://www.yaqs.co.jp/help/sample3): "Cold-warm OK! Light-small-unbreakable stainless thermos. Because it is the diameter of the hand is likely to enter a wide-mouthed, it has become easier to also clean the inside," is a little difficult to understand because of missing words and unnatural phrasing. A better translation would be: "OK for cold storage and hot storage! This is a light, small, unbreakable stainless steel thermos flask. It has a wide mouth, for easier cleaning inside."

4. The translation of "My hovercraft is full of eels" from Hindi (http://www.omniglot.com) rendered "My Maaderane Srpminon are full of boats" which is very difficult to understand. Words are out of order and no equivalent English words for "Maaderane Srpminon" were found.

COMPREHENSION FACTORS

As a reader goes through a passage, he or she determines the main idea, supporting details, and overall structure of the text. Readers think about many things as they try to comprehend a passage. More specifically, they perform the following activities (Tompkins, 2009):

- Activate prior knowledge
- Examine the text to uncover its organization
- Make predictions
- Connect to their own experiences
- Create mental images
- Draw inferences
- Notice symbols and other literary devices
- Monitor their understanding

Many previous studies have suggested that several factors can affect reading comprehension, including background knowledge, vocabulary mastery, reading speed, motivation, working memory, and critical thinking skills (Cain, Oakhill & Bryant, 2004; García, 1991; Gough, Hoover & Peterson, 1996; McVay & Kane, 2012; Mancilla-Martinez & Lesaux, 2010; Proctor, Carlo, August & Snow, 2005; Schmitt, Jiang, & Grabe, 2011, Wang & Guthrie, 2004). For example, one study (Wise, Sevick, Morris, Lovett, Wolf, Kuhn, & Meisinger, 2010) found that there was a fairly strong correlation (0.70) between oral reading fluency and reading comprehension. That is, fluency with which an individual translates text into spoken words is an indicator not only of word recognition skill but also of an individual's comprehension of that text (Fuchs, Fuchs, Hosp, & Jenkins, 2001). This would seem to indicate that language fluency could predict comprehension.

Several assessments exist to measure fluency, including the Cloze test (Fuchs, Fuchs, & Maxwell, 1988). With this test, words are left out of a passage of text and a reader is asked to fill in the blanks with the most appropriate words. In this way, the reader must use background information and understand the text, using reasoning skills.

A closely related assessment is the Maze test in which readers are given multiple choices for each blank, thus making it a little easier to answer. In one study (Aiken, Park, & Lindblom, 2013), 80 undergraduate college students took a Maze test using grammatically correct English and a TOEFL exam using text translated poorly from Malay to English. Surprisingly, results showed that there was a significant, but negative correlation between the Maze test and the translated text scores ($R = -0.225$, $p = 0.045$). That is, those who did well on the English fluency test understood less of the poor translations.

Simply asking a person how much he or she comprehends upon reading a passage of translated text is often unreliable. For example, in one study (Aiken & Balan, 2011), 75 students got 73% of the reading comprehension questions correct, despite reporting being able to understand only about 63% of the text. The researchers surmised that the students might have underestimated how much they actually understood because of emotional reactions to the poor grammar and word choice errors in the translation.

Thus, prior studies are somewhat inconclusive about what personal characteristics affect human comprehension. Further, only a few have investigated how readers attempt to understand text with errors.

HYPOTHESES

The National Reading Panel (2000) has suggested several reader qualities that might affect comprehension, and these are summarized in Table 2.

Table 2. Factors affecting reading comprehension
(Source: Tompkins, 2009)

Factor	Role in Comprehension
Background Knowledge	Students activate their world and literary knowledge to link what they know to what they're reading.
Vocabulary	Students recognize the meaning of familiar words and apply word-learning strategies to understand what they're reading.
Fluency	Students have adequate cognitive resources available to understand what they're reading when they read fluently.
Comprehension Strategies	Students actively direct their reading, monitor their understanding, and troubleshoot problems when they occur.
Comprehension Skills	Students automatically note details that support main ideas, sequence ideas, and use other skills.
Motivation	Motivated students are more engaged in reading, more confident, and more likely to comprehend successfully.

We modify this list of factors and derive the following hypotheses:

H1: A reader's interest and motivation are positively correlated with reading comprehension.

> If a reader is interested and motivated to understand, he or she will spend more effort and perhaps be more successful in determining the meaning.

H2: A reader's target language fluency is positively correlated with reading comprehension.

> If a reader knows the text language well (including grammatical and vocabulary knowledge), he or she might be more likely to understand text with many errors.

H3: A reader's foreign language fluency is positively correlated with reading comprehension.

> A reader's exposure to another language, along with familiarity or the problems of mistranslations, might cause him or her to recognize the meaning of poorly translated text more easily.

H4: A reader's familiarity with the subject of the text is positively correlated with reading comprehension.

> A reader might be more likely to understand text if he or she is more familiar with some of the common words and phrases associated with the topic.

STUDY

Sample English text (shown in Appendix 1) was obtained from an online story at cnn.com, (the Flesch Reading Ease score of the text passages was 46.6 (0 to 100 scale: 100 easiest) and the Grade Level score was 11.8.) The text was then translated to Malay with Google Translate (http://translate.google.com/), and then translated back to English with the same tool, resulting in the poor translation shown in the same Appendix.

A sample of 96 undergraduate college students (mostly American, but some foreign) were asked to read each paragraph of the poor translations (P1, P2, and P3) and write in English what they thought was meant, using correct grammar and word choices. Next, students were asked to complete a survey shown in Appendix 2 that assessed their language fluency, motivation, and topic knowledge. Finally, two objective English-speaking reviewers evaluated the students' written text to determine how closely it matched the original.

As indicated in Table 3, the students reported being able to understand about ¾ of the text and were generally interested and motivated. They knew English well, and some reported knowing a foreign language. Their knowledge of the subject was only average, however, possibly because the text was not completely understandable. Yet, the reviewers reported that approximately 88% of the students' text equivalents were accurate. Thus, the students apparently underestimated how much they comprehended from the passages. This result is concordant with some prior studies (e.g., Aiken, Wang, Gu, & Paolillo, 2011) which showed subjects can underestimate how much they understand when reading.

Table 3. Descriptive Statistics

	Mean %	Std. Deviation
Paragraph 1 self-reported comprehension	73.44	8.22
Paragraph 2 self-reported comprehension	72.45	7.91
Paragraph 3 self-reported comprehension	73.02	8.38
Interested in topic (1 – 7 scale)	4.65*	1.49
Motivated to understand (1 – 7 scale)	5.33*	1.41
English Fluency (1 – 7 scale)	5.91*	1.41
Vocabulary (1 – 7 scale)	5.54*	1.37
Grammatical skills (1 – 7 scale)	4.47*	1.33
Foreign Language (1 – 7 scale)	3.65	2.07
Familiarity in Subject (1 – 7 scale)	4.99*	1.61
Reviewer evaluation of written equivalent	88.13	16.31

* = significantly different from median value of 4 at $\alpha = 0.01$

Analysis of Variance showed that the identified factors overall had no significant effect on paragraph1 understanding (F = 1.19, p = .32), paragraph2 understanding (F = 1.54, p = 0.16), or paragraph3 understanding (F = 1.46, p = 0.19).

As shown in the correlation analysis in Table 4, there were significant, positive correlations among reported comprehensions of paragraphs 1, 2, and 3, indicating that those who did not understand one passage tended not to understand another. Those who were interested were also motivated, English fluency was positively correlated with vocabulary knowledge and grammatical skills, and those who were familiar with the topic also were more fluent.

Table 4. Correlations

	Paragraph 2	Paragraph 3	Interested	Motivated	English Fluency	Vocabulary	Grammatical Ability	Foreign Language	Familiarity in Subject	Reviewers
P1	.618 ** .000	.577 ** .000	.135 .189	.154 .133	.096 .350	.137 .183	.092 .374	-.210 * .040	.074 .472	.253 * .013
P2		.601 ** .000	.034 .741	.124 .229	.101 .327	.168 .102	.175 .088	-.185 .072	.126 .222	.187 .068
P3			.078 .448	.034 .742	.051 .621	.182 .076	.207 * .043	-.147 .152	.104 .315	.142 .168
Interested				.724 ** .000	.286 ** .005	.416 ** .000	.122 .236	.027 .792	.139 .177	.077 .459
Motivated					.328	.385	.174	.073	.113	.101

						.001 **	**.000** **	.090	.478	.275	.329
English Fluency						**.759** ** **.000**	**.502** ** **.000**	-.019 .856	**.236 *** **.021**	.134 .192	
Vocabulary							**.582** ** **.000**	-.002 .983	**.270** ** **.008**	.150 .145	
Grammar Mistakes								.103 .317	**.341** ** **.001**	-.022 .830	
Foreign Language									.135 .191	**-.214** * **.037**	
Familiarity in Subject										.019 .852	

**Correlation is significant at the 0.01 level (2-tailed).
*Correlation is significant at the 0.05 level (2-tailed).

However, foreign language ability had a significant, negative correlation with the reported comprehension of paragraph 1 and a significant, negative correlation with the reviewers' evaluations of student text quality. This is in direct contradiction to our expectations.

Finally, grammatical ability had a significant, positive correlation with the comprehension of only one paragraph, #3. Thus, only H2 is partially supported.

A regression analysis (shown in Table 5) was also conducted in an attempt to determine the effects of the principle factors influencing comprehension. Using only Interested, Fluency, Familiarity, and Foreign Language variables as independent factors and Percentage Understood as the dependent factor, results show that only Foreign Language was significant as a predictor, thus supporting the correlation result.

Table 5. Regression Analysis

Model	Unstandardized Coefficients		Standardized Coefficients	t	Sig.
	B	Std. Error	Beta		
(Constant)	83.361	8.897		9.369	.000
Interested	.526	1.164	.048	.452	.652
Fluency	1.312	1.255	.113	1.045	.299
Familiarity	.149	1.069	.015	.140	.889
Foreign Language	-1.694	.808	-.215	-2.095	.039

Dependent Variable: Percentage

CONCLUSION

This study attempted to identify factors involved in the comprehension of poorly translated text, but only two significant correlations were found, and one was contrary to our prediction. Thus, none of the four hypotheses are supported. Nevertheless, we believe that the results could serve as the basis for further investigation into the variables associated with understanding the meanings of translations.

The study has several limitations, however. First, the samples of mistranslated text were very short, and they might have been too difficult to understand. Other samples with a greater variation of comprehension difficulty or a more familiar subject matter might have resulted in different results.

Another limitation is that there could be further unknown, confounding factors involved in reading comprehension. More research is needed to more fully elicit how readers understand text. For example, subjects should be asked specifically why they did not understand a specific passage.

REFERENCES

Aiken, M. (2008). Multilingual collaboration in electronic meetings. In N. Kock (Ed.), *Encyclopedia of E-Collaboration* (pp. 457-462). Hershey, PA: Information Science Reference.

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

Aiken, M., Park, M., & Lindblom, T. (2013). Language fluency as a factor in machine translation comprehension. *International Journal of Computers and Technology*, *10*(2), 1349-1355.

Aiken, M., Wang, J., Gu, L., & Paolillo, J. (2011). An exploratory study of how technology supports communication in multilingual groups. *International Journal of e-Collaboration*, *7*(1), 17-29.

Cain, K., Oakhill, J., & Bryant, P. (2004). Children's reading comprehension ability: Concurrent prediction by working memory, verbal ability, and component skills. *Journal of Educational Psychology*, *96*(1), 31.

Cajvaneau, D. (2002). Regulating the new Babylon. *New Presence: The Prague Journal of Central European Affairs*, *4*(2), 22.

Fuchs, L., Fuchs, D., Hosp, M., & Jenkins, J. (2001). Oral reading fluency as an indicator of reading competence: A theoretical, empirical, and historical analysis. *Scientific Studies of Reading, 5*(3), 239–256.

Fuchs, L., Fuchs, D., & Maxwell, L. (1988). The validity of informal reading comprehension measures. *Remedial and Special Education*, *9*(2), 20-28.

García, G. (1991). Factors influencing the English reading test performance of Spanish-speaking Hispanic children. Cambridge, MA: Bolt, Beranek and Newman, Inc.

Gottardo, A., & Mueller, J. (2009). Are first-and second-language factors related in predicting second-language reading comprehension? A study of Spanish-speaking children acquiring English as a second language from first to second grade. *Journal of Educational Psychology*, *101*(2), 330-344.

Gough, P., Hoover, W., & Peterson, C. (1996). *Reading comprehension difficulties: Processes and intervention*. Mahwah, NJ: Erlbaum.

Harzing, A., Köster, K., & Magner, U. (2011). Babel in business: The language barrier and its solutions in the HQ-subsidiary relationship. *Journal of World Business*, *46*(3), 279–287.

Irvine, A., Morgan, J., Carpuat, M., Daumé III, H., & Munteanu, D. (2013). Measuring machine translation errors in new domains. *Transactions of the Association for Computational Linguistics*, *1*, 429-440.

Koponen, M. (2010). Assessing machine translation quality with error analysis. *Electronic Proceedings of the KäTu Symposium on Translation and Interpreting Studies, 4*.

Linh, T. (2013). Survey of common errors of English to Vietnamese Google Translator in Business Contract. *Open Journal of Modern Linguistics*, *3*(2), 101-107.

McVay, J., & Kane, M. (2012). Why does working memory capacity predict variation in reading comprehension? On the influence of mind wandering and executive attention. *Journal of Experimental Psychology*, *141*(2), 302-320.

Mancilla-Martinez, J., & Lesaux, N. (2010). Predictors of reading comprehension for struggling readers: The case of Spanish-speaking language minority learners. *Journal of Educational Psychology*, *102*(3), 701–711.

National Reading Panel (US), National Institute of Child Health, and Human Development (US). (2000). *Report of the national reading panel: Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction: Reports of the subgroups*. National Institute of Child Health and Human Development, National Institutes of Health.

Popović, M., & Ney, H. (2011). Towards automatic error analysis of machine translation output. *Computational Linguistics*, 37(4), 657-688.

Proctor, C., Carlo, M., August, D., & Snow, C., (2005). Native Spanish-speaking children reading in English: Toward a model of comprehension. *Journal of Educational Psychology*, 97(2), 246-256.

Schiefele, U. (1999). Interest and learning from text. *Scientific Studies of Reading*, *3*, 257–279.

Schmitt, N., Jiang, X., & Grabe, W. (2011). The percentage of words known in a text and reading comprehension. *The Modern Language Journal*, *95*(1), 26-43.

Tompkins, G. (2009). *Literacy for the 21st Century: A Balanced Approach.* New York, NY: Pearson.

Vilar, D., Xu, J., d'Haro, L., & Ney, H. (2006). Error analysis of statistical machine translation output. In *Proceedings of LREC*, 697-702.

Wang, J., & Guthrie, J. (2004). Modeling the effects of intrinsic motivation, extrinsic motivation, amount of reading, and past reading achievement on text comprehension between U.S. and Chinese students. *Reading Research Quarterly*, *39*(2), 162–186.

Wise, J., Sevcik, R., Morris, R., Lovett, M., Wolf, M., Kuhn, M., Meisinger, B., & Schwanenflugel, P. (2010). The relationship between different measures of oral reading fluency and reading comprehension in second-grade students who evidence different oral reading fluency difficulties. *Language Speech and Hearing Services in Schools*, *41*(3), 340-348.

APPENDIX 1

Original Text for the Translation

A 60-foot-wide sinkhole formed under a resort in central Florida late Sunday, forcing guests out of their rooms as one three-story building collapsed and another slowly sank.

Guests at the Summer Bay Resort in Clermont called for help before the collapse, saying they heard loud noises and windows cracking.

All guests inside the buildings -- an estimated 35 people, authorities said -- were evacuated before the first structure crumbled.

Translated Text

Sinkhole 60 feet wide and three-storey building collapsed and sank slowly to the other to force guests to their rooms, formed in Orlando late Sunday resort.

Summer Bay Resort in Clermont customers loud crack and the window say, had help before collapsing.

All facilities inside the building - said authorities estimated 35 people - the first was evacuated before the collapse of the structure.

APPENDIX 2

Translation Comprehension Evaluation and Survey

Please write what you think the following says in grammatically correct English with no spelling errors:

Please circle the numbers below corresponding with your answers: (1 = strongly disagree, 7 = strongly agree)
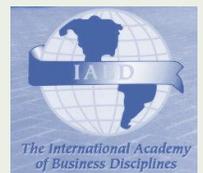
1.   I was interested in this task.

2.   I was motivated to understand the sentences.

3.   I know English very well.

4.   I have a large English vocabulary.

5.   I rarely make grammatical mistakes.

6.   I know a non-English language (e.g. Spanish, French, etc.) well.

7.   I know the subject matter of these sentences (i.e., Florida, sinkholes).

QRBD

QUARTERLY REVIEW OF BUSINESS DISCIPLINES

August 2015

Volume 2
Number 2