

Reading Comprehension and the Use of Google Translator

Adriana Riess Karnal¹ & Pereira W. Vera¹

¹ Languages Department, Universidade do Vale do Rio dos Sinos, Brazil

Correspondence: Adriana Riess Karnal, Centro de Ciências da Comunicação, Unisinos University, São Leopoldo, Brazil. Tel: 51-591-1100. E-mail: adrikarnal@gmail.com

Received: December 19, 2012 Accepted: September 19, 2013 Online Published: November 24, 2013

doi:10.5539/ijel.v3n6p113

URL: <http://dx.doi.org/10.5539/ijel.v3n6p113>

Abstract

This paper analyses the use of Machine Translation when EFL students read a text translated by Google translator. Bearing this in mind we examine how efficient Google translator can be and to what extent it is possible to comprehend a text translated with this tool. We discuss the area of Reading through a psycholinguistic point of view and we focus mainly on Perfetti's Verbal Efficiency (1985, 2007). Finally, we applied a Reading instrument to a group of non-proficient readers of English. The objective is to investigate the strategies readers use when reading a text translated from English to Portuguese and how much they could comprehend.

Keywords: reading comprehension, machine translation, reading strategies

1. Introduction

Regarding linguistics research that orients reading in English as a foreign language, it is not possible to deny how Google Translator has had an impact to the readers. This is because studies in Machine translation (Hutchins, 2003) have been developed, and Google translator is an example of this development. It can translate texts so that we can partly comprehend. This paper has the objective to analyze reading comprehension when learners of English as an L2/FL read a text translated from English into Portuguese using Google translator. The theoretical background presented here is mostly based on Charles Perfetti's verbal efficiency model (1985, 2001, 2007).

This article is divided in four sessions. The first one introduces Google electronic translator and its performance when translating from English into Portuguese. The second session reviews researches about reading and we analyze Perfetti's model entitled Verbal efficiency. Next, the third session is more practical as it describes a research conducted with a group of basic level undergraduate students of English. A reading instrument (reading test) was run in order to check the level of comprehension. Finally, the last session discusses the results of the instrument.

1.1 Google Translator

One of the *foci* of computational linguistics is the development of electronic translators, an area called *Machine Translation* (Hutchins, 2003). As for Google translator, it is grounded on a process entitled Statistical Machine Translation (henceforth SMT) which searches patterns in great amount of texts. Initially, translators worked on the lexical level that was not different from dictionaries. As long as researches progressed, electronic translators came to work on a textual or discursive level, as they could carry out verbal dependency, collocations and contextual recognition. According to Lima (2008):

Recent researches using bilingual and multilingual database have improved as they recognize textual typology (genre) and grammar rules, such as government, collocations, word order, polysemy, idioms, as well as attach dictionaries to the database in order to translate technical texts (p. 3)

It is visible the progress of Google translator, it translates today to 57 languages. To exemplify, two texts were electronically translated from Chinese and English into Portuguese. The first text refers to an interview by the Brazilian diplomat Paulo Roberto de Almeida published on the Chinese newspaper *Shanghai Wenhui Daily* in November 2011. The second one refers to an article on *The New Yorker* about Brazilian president Dilma Rousseff.

Text 1)

环球视窗

智利学生抗议浪潮愈演愈烈 邻国开始反思自身制度弊端

南美国家家家有本难念的教育经

日期: 2011-10-21 作者: 朱幸福 来源: 文汇报

图片作者:

图片作者:

图片作者:

图片作者:

本报首席记者 朱幸福

(...) (本报巴西利亚 10 月 19 日专电)

Google Translation:

window global

Onda de protestos estudantis no Chile intensificou os vizinhos começaram a refletir sobre os seus defeitos próprio sistema

Países sul-americanos, cada família tem sido difícil de estudar educação

Data: 2011 Autor -10-21: Zhu Fonte felicidade: Wen Wei Po

Foto por:

Foto por:

Foto por:

Foto por:

Chefe do jornal correspondente a felicidade Zhu

(...)(Brasília jornal Ta Kung Pao, 19 de Outubro)

Text 2)

ABSTRACT: A REPORTER AT LARGE about Brazilian President Dilma Rousseff. Until recently, Brazil has been one of the most uneducated, economically imbalanced countries in the world. Now its economy is growing much more rapidly than that of the U.S. Twenty-eight million Brazilians have moved out of severe poverty in the past decade. The country has a balanced budget, low national debt, nearly full employment, and low inflation. It is, chaotically, democratic, and it has a free press. Brazil operates in ways we have been conditioned to think are incompatible with a successful free society. It isn't just that Brazil is ruled by unapologetic former revolutionaries, many of whom—including the President—were imprisoned for years for being terrorists. The central government is far more powerful and intrusive than it is in the U.S. It is also far more corrupt. Crime is high, schools are weak, roads are bad, and ports barely function. And yet, among the world's major economic powers, Brazil has achieved a rare trifecta: high growth, political freedom, and falling inequality.

Google Translation:

RESUMO: A REPORTER em geral sobre o presidente brasileiro, Dilma Rousseff. Até recentemente, o Brasil tem sido um dos mais ignorantes, os países economicamente desequilibrado do mundo. Agora a economia está crescendo muito mais rapidamente do que os EUA Vinte e oito milhões de brasileiros saíram da pobreza extrema na última década. O país tem um orçamento equilibrado, a dívida nacional baixa, quase o pleno emprego e inflação baixa. É, caoticamente, democrático, e tem uma imprensa livre. Brasil opera de maneiras que foram condicionados a pensar são incompatíveis com uma sociedade livre bem sucedida. Não é apenas que o Brasil é governado por ex-revolucionários sem remorso, muitos dos quais, incluindo o presidente, foram presos por anos para serem terroristas. O governo central é muito mais poderosa e intrusiva do que em os EUA Também é muito mais corrupto. Criminalidade é alta, as escolas são fracas, as estradas são ruins, e as portas mal função. E ainda, entre os grandes poderes econômicos, o Brasil alcançou um trifecta raros: alto crescimento, a liberdade política e a desigualdade em queda.

Both texts demonstrate the reader is able to have a certain level of reading comprehension. Although the Chinese translation is more difficult, because it looks like a telegraphic message, we can depict the main idea (skimming

reading), that is the protests in Chile. Linguistics subtleties, such as referring pronouns (anaphora) seen in the pronoun “seus” is ambiguous, since it is not possible to know if it refers to “vizinhos” or “sistema”. The sentence that follows “*cada familia tem sido dificil ...*”, even the verb is not adequate we can infer the families have had difficulties with education. Anyway, Google translator gave lexical access once the Chinese ideograms do not look like the Latin characters. It is important to highlight that studies about *Reading* believe lexical understanding is fundamental to comprehension (Perfetti, 2007; Tomitch, 2009; Kamil e Hiebert, 2005).

The translation in the second text is more qualified. It is possible to understand the text in general terms, as we can scan the text for the details, in a way we could perceive subtleties concerning what is said about the government opposition. It is worthwhile to mention the word “trifecta”, both in English and in Portuguese it is the same word. However, it is not an ordinary word, it is a borrowing from the Turf. Thus, it can represent some difficulty to the reader, although the prefix *tri* points to the number *three* and this is what happens when the author cites these three items- high growth, political freedom, and falling inequality. In this case, morphology is the source of information to the reader. Perfetti (2007) and Kratochvil (2010, p. 68) explain that the analyses of morphological characteristics of a word is a strategy to be used. On the other hand, Juell (1999) states that good readers do not use context to infer meanings, once they know the words. What has to be taken into consideration is that being aware of morphology is a way of knowing a word.

In fact, both texts are likely to be read and can be comprehended. This leads to a positive performance from Google translator. In order to analyze reading comprehension in the English/Portuguese article above, a questionnaire was run to a group of Brazilian undergraduates who are not proficient in English. Before presenting this analysis we support our ideas based on Perfetti’s model described below.

2. Theoretical Framework on Reading and Cognition

Literature about Reading has dealt with two processes called *top down and bottom up*. In the first process the reader puts into practice his or her own linguistic knowledge, such as vocabulary, prefixes, suffixes, textual/discursive markers, textual organization, conjunctions, etc. In the second one, the reader interacts with the text and cognitive mechanisms are at stake, such as memory, inference and predictability. Perfetti (2001) says these cognitive mechanisms assemble words into messages, that is why they are vital to explain how Reading occurs. Many authors today explain both processes work together (Perfetti, 2001; Eskey, 1988; Grabe, 1988; Smith, 1999) interactively.

Yet, world knowledge, also described as Rumelhart schemata’s theory, is of great importance to Reading comprehension. Then, knowing the genre, who the author is, as well as grammatical implicit knowledge are all knowledge we assume before reading itself. The theory says bottom up processing triggers cognitive mechanisms involved in reading. Rumelhart explains (1980):

Schemata can represent knowledge at all levels—from ideologies and cultural truths to knowledge about the meaning of a particular word, to knowledge about what patterns of excitations are associated with what letters of the alphabet. We have schemata to represent all levels of our experience, at all levels of abstraction. Finally, **our** schemata are our knowledge. All of our generic knowledge is embedded in schemata.

The scheme the author talks about represents our experience in the world and is responsible for general knowledge we acquire. He adds “Schemata are organized mental structures that allow the learners to understand and associate what is being presented to them”. In addition to represent our knowledge, schemata are at the same time organizers of what we know. Then, they are likewise in memory. Rumelhart says that knowledge is disassembled into categories and stored into memory. We might say *Schemata Theory* is a theoretical model which aims at representing external world in an internal world. Similar to Rumelhart’s is Fillmore’s frame theory. However, Fillmore believes his theory is more culturally oriented as it refers to the community of speakers.

We shall not forget that knowledge is stored in memory; therefore it is also a main tenet for reading. Numminen (2002, p. 1) states that reading has straight relations to memory, for example, an insufficient working memory capacity, or a disorganized long term memory can cause difficulties to comprehension. It is worth reminding about these two memories just mentioned—working memory and long term memory. The first means cognitive processes keep information in mind during active information processing. It enables decoding and information register here and now. Numminen (2002, p. 2) goes further: “*Working memory is a especially significant memory area for learning to read.*” The second one is an unlimited storage which retrieves and organizes all we know and do. Therefore, world knowledge is also organized by long term memory, and it is, according to what we described previously, indispensable to read.

In Perfetti's efficiency model, limitations of working memory and lexical processes are the tenets of Reading comprehension. Afterwards, along with Marron e Folz the author will add other principles to his theory, quoting Baldo (2006, p. 33): ... inferences, comprehension monitoring, knowledge of content and word meaning should be considered if the final objective was a wide explanation to Reading processing.

Perfetti's theory also talks about inferences and lexical access. He is, in fact, describing cognitive strategies we use while reading. As for monitoring comprehension, it is a metacognitive strategy, that is, the one we can control consciously. When we think about an efficient methodology for teaching Reading in English the use of these strategies must be present. For Tomitch (2009, p. 192), the focus of the lesson must be on Reading comprehension not on grammar and vocabulary solely. It is in this regard we must equip learners, thus, teaching strategies is the tool.

3. Methods: Application of a Reading Instrument

Most universities today offer Reading English lessons in their curricula. As a methodological approach teachers all agree that the development of cognitive strategies for learning Reading in English as an L2 is more accurate than the old design of teaching grammar or vocabulary solely. For this reason, this paper seeks to analyze Reading comprehension when facing a text translated by Google and what strategies were used during this task. We believe this study can better explain how a non-proficient reader deals with this kind of translation, because the machine cannot master a translation as a human would do.

One of the central ideas is that the non-proficient reader does not have lexical, grammar or cultural access to the knowledge implied in the text. That would make reading impossible, such as the text presented here in Chinese, for example. Google translation makes, at least, words accessible. It is the learner's effort to use strategies to better comprehend what he/she is reading.

From what has been said so far, we present next the results of a reading instrument applied to a group of non-proficient undergraduate students. The subjects should read a text (the article about Brazilian president in section 1) and answer five questions in which reading comprehension was assessed. Table 1 shows the profile of the subjects.

Table 1. Profile of the readers

| Description of the subjects | |
|--------------------------------------------------|-----------------|
| Number of subjects | 8 |
| Level of English as an FL | Basic |
| Education | undergraduation |
| Age | 18-40 |
| Number of questions to be answered in Portuguese | 5 |

The results of the correct answers are shown in table 2 below.

Table 2. Correct answers

| Question nb | Nb. of subjects | Results in % |
|----------------------------|-----------------|--------------|
| 1 | 6 | 81,25 % |
| 2 | 4 | 56,25 % |
| 3 | 3 | 56,25 % |
| 4 | 5 | 68,75 % |
| 5 | 4 | 50 % |
| Average of correct answers | | 62,5 % |

4. Discussion

We believe the positive result of the first question is due to the quality of linguistic processing while students

read. The translation was very clear where the reader had to scan the text to find the correct answer. At the sentence level there were no problems of word order, pro-dropping, or any syntactic rule that could misinterpret reading. It seems that while readers scanned the article to search for the right answer reading flowed. For Perfetti (2001), fluency is related to comprehension.

Concerning the second question the results are below average. This is because the answer implied inferences. The reader not only had to scan the text from lines 7 to 13 to find the correct answer, as they also had to predict and use world knowledge. Moreover, there was a great deal of syntactic errors in the translation. Therefore, parsing the sentences might have been difficult, causing difficulties to store information in working memory.

As for the third answer the result is not positive either. The problem of comprehension here might be in the semantic-pragmatic level. The word *terrorist* is a reference to the Brazilian president in the text, however this is not a typical reference to any president. Readers should count on their world knowledge (Rumelhart's schema) to answer this question. In addition to that, the words *terrorist* and *president* are separated by a relative pronoun and are distant from each other. Parsing might have been difficult; again, it could have overloaded working memory.

The result of the fourth question is positive. Readers scanned the text correctly to find this answer. It is probably because there are no problems in terms of syntax at the sentence level. It seems the only trouble is related to the word *intrusive* which refers to the central government. This can be considered a rare word in the Portuguese vocabulary. Perfetti (2007) explains that rare words do interfere in reading comprehension, simply because readers do not know them, and sometimes it is impossible to infer from the context.

The last question has 50 % of correct answers and 50% of no answers at all. Again, there is a problem at the word level due to the use of the rare word *trifecta*. The ones who answered correctly must have used morphology (prefix *tri*) to infer the meaning and contextualize the word. The ones who did not know the word at all gave up answering. It might have happened because that was the last answer and readers might have been mentally stressed out.

5. Conclusion

This paper had the main objective to analyze Reading comprehension by readers of English as an SL/FL of a text translated with Google translator. As a general result readers could partially understand the article that was translated. In fact, the result of 62% of comprehension is more than the half of the text, and it can be considered positive if we bear in mind the readers are not proficient. The difficulties were due to linguistic or cultural inability of the readers, as Perfetti (1985, 2007) has stated.

The result showed us that it is possible to use machine translation to both partially comprehend texts and use it to enhance knowledge in the English Reading lesson. That means Google translator can work as didactic material, that is, the translated text can work as linguistic input students must be attentive. What we mean is that the teacher who uses electronic translation in class can develop strategies so as to orient the reader to the errors the machine made or not.

Finally, we would like to discuss how Google translator has been developed and how much the reader and teachers can use it as a tool to teach or learn. We should teach our students to be independent of any translator while they are reading; however, electronic tools can be helpful to make students linguistically aware while they are in the process of learning.

In the future we will probably see more accurate electronic translators. In the area of SLA/FL it is a need to understand the role of teaching reading when this tomorrow comes.

References

- Alderson, C. J. (2000). *Assessing reading*. Cambridge. Cambridge University Press. <http://dx.doi.org/10.1017/CBO9780511732935>
- Anderson, N. J. (1991). Individual differences in strategy use in second language reading and testing. *The Modern Language Journal*. <http://dx.doi.org/10.1111/j.1540-4781.1991.tb05384.x>
- Baldo, A. (2006). *O uso de estratégias de língua materna e na língua estrangeira*. Tese de doutorado. Porto Alegre. PUC RS.
- Eskey, D. (1986). Theoretical foundations. In *Teaching Second Language Reading for Academic Purposes*. Massachusetts: Addison Wesley Publishing Company.
- Fillmore, C. (1985a). *Frames and the semantics of understanding*. Quaderni di Semantica.

- Finger kratochvíl, C. (2010). *Estratégias para o desenvolvimento da competência lexical: Relações com a compreensão leitora*. Tese de Doutorado. Florianópolis. Ed. UFSC.
- Hutchins, J. (2003). Machine translation: General overview. In *The Oxford Handbook of Computational Linguistics*. Oxford University Press.
- Juel, C. (1999). *Learning to read words* (1st ed.). University of Michigan Press.
- Kamil, M., & Hiebert. (2005). *Teaching and learning vocabulary: Bringing research to practice*. Mahwah, NJ. Erlbaum.
- Klein, A. (2008). *Oficina de leitura: Teoria e prática*. 12a. São Paulo. Ed. Pontes.
- Lima, L. (2011). Tradução eletrônica: do riso irônico ao interesse científico. *Fólio Revista de Letras*, 3(1).
- Numminen, H. (2002). *Working memory in adults with intellectual disability*. Helsinki. Famr Research Publications.
- Nunan, D. (1991). *Language teaching methodology: A textbook for teachers*. Prentice Hall.
- Pereira, V. W. (2005/2007). *Predição leitora em ambiente virtual e ambiente não-virtual: Ensino, pesquisa e extensão*. In: Prele3.
- Pereira, W. V. (2001/2002). *Aprendizado da leitura e da escrita através do computador por alunos do ensino fundamental*. In: ALEC.
- Perfetti, C. A. (1985). *Reading ability*. New York: Oxford University Press
- Perfetti, C. A. (1999). Comprehending written language: A blueprint of the read. In M. C. Brown, & P. Hagoort (Eds.), *The neurocognition of language*. Oxford: OUP.
- Perfetti, C. A. (2007). *Reading Ability: Lexical quality to comprehension*. Lawrence Erlbaum.
- Perfetti, C. A. Marron M.F. (1996) Sources of comprehension failure: theoretical perspectives and case studies. In *Reading Comprehension Difficulties*. NJ, Erlbaum.
- Rumelhart, D. E. (1980). Schemata: The building blocks of cognition. In R. J. Spiro et al. (Eds.), *Theoretical issues in reading comprehension*. Hillsdale, NJ. Lawrence Erlbaum.
- Smith, F. (1999). *Leitura significativa*. Porto Alegre. Ed. Art Med.
- Stanovich, K. E. (1980). Towards an interactive compensatory model of individual differences in the development of reading fluency. *Reading Research Quarterly*, 16. <http://dx.doi.org/10.2307/747348>
- Tomitch, L. (2009). Aquisição da leitura em língua inglesa. In *Ensino e aprendizagem de língua inglesa: conversas com especialistas*. Org. Lima, Diógenes. Ed. Parábola.

Copyrights

Copyright for this article is retained by the author(s), with first publication rights granted to the journal.

This is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/3.0/>).