



Reading Strategies in a L2: A Study on Machine Translation

Adriana Riess Karnal

Pontifícia Universidade Católica do Rio Grande do Sul PUC-RS

Vera Vanmacher Pereira

Pontifícia Universidade Católica do Rio Grande do Sul PUC-RS

ABSTRACT

This article aims at understanding cognitive strategies which are involved in reading academic texts in English as a L2/FL. Specifically, we focus on reading comprehension when a text is read either using Google translator or not. From this perspective we must consider the reading process in its complexity not only as a decoding process. We follow the paradigm of psycholinguistics in the reading area and we also describe reading strategies classified by some authors as global or local and cognitive or metacognitive. In order to compare how reading strategies are used when readers either utilize Google translator or not 10 students from the Language Institute from the University of Pittsburgh were tested. The methodology *Think aloud protocols* was used so that we could compare and analyze if machine translation was a benefit to reading comprehension.

Keywords: reading strategies, reading comprehension, machine translation

INTRODUCTION

This article aims at understanding strategies involved in reading academic texts in English as a L2 /FL. Specifically, we focus on reading comprehension when a text is read either utilizing Google translator or not. Some circumstances contribute to this research. First, the importance of cognitive strategies and reading, a relatively recent research that has proven to be effective to explain the reading process. Also, we briefly describe the area of Machine translation, since the use of Google translator is used as a support strategy. This investigation analyses the possibilities of using Google translator as a support tool for reading in a L2. The research was conducted based on data collected at the English Language Institute (ELI) at the University of Pittsburgh. Ten participants participated and all of them were speakers of other mother languages other than English.

The study is divided into four sections. The first one deals with the concept of reading comprehension and its cognitive underpinnings. In the second section we discuss the role cognitive strategies play while one is reading. What we want to understand is how these strategies are used when the reader either uses Google translator or not. Furthermore, we want to know to what extent a poor translation is capable of enhancing understanding.

The third section of this paper presents how the instruments were conducted, so it has a focus on the methodology used. The goal is to evaluate how those students whose L1 is not English

understand a text read either with Google or without it. Finally, the last section concludes the work analyzing data collected.

BACKGROUND

1.1 What is reading comprehension?

In order to understand what reading comprehension is, it is necessary to emphasize that the purpose of reading varies among readers. According to Koda (2005), we can read to learn scientific content, to get some information in a list or to find entertainment. Whatever the initial motivation of reading is, the ultimate goal is always comprehension. Also, it is important to say that there are, at least, two levels of understanding. The first level is shallow, the reader can find explicit information in the text. The other level is deeper, it depends on inferential abilities of the reader, because textual information is not always explicit. This level of comprehension only occurs when the reader combines prior knowledge they have, that is, their culturally constructed knowledge stored in a mental scheme as well as the inferential ability.

In general terms, reading comprehension is the process of simultaneously extract and construct meaning through involvement with written language. Caldwell (2008) explains that comprehension includes three parts: a) the active process of understanding that comes from an ability (or strategy), b) world knowledge and c) motivation to read a text. Kintsch and Kintsch (2005) say that comprehension is not a single process, the reader articulates a variety of concurrent processes, such as identification of words, for example through phonology and comprehension processes. Reading comprehension occurs when the reader means the text, and it does not occur by summing up the meanings of individual sentences, on the contrary, it is based on the implicit or explicit coherence relations in the text which are processed in the mind of the reader through inferential activity. In order to achieve this it is necessary to perform such cognitive operations of logical connections to world knowledge.

We have to bear in mind that both inference and world knowledge are present in reading comprehension. This suggests that comprehension processing is not equal from reader to reader, since there are individual differences in the inferential ability and knowledge or experience acquired over a lifetime. Thus, the cognitive and cultural issues are embedded in the use of reading strategies. Inference itself is a cognitive ability and monitoring is a metacognitive one. For instance,

using a dictionary to monitor what is not known is an ability the reader is aware of, it is deliberate, so is machine translation. The question that arises is: What are the reading strategies used with and without the use of Google translator? What is the relationship between such strategies and reading comprehension?

Regarding Google translator, because of globalization and easy access to the internet, this tool has been increasingly demanded socially. Nevertheless, having proficiency in a foreign language is a long process, for this reason, the use of Google translator is seen as an aid to understand texts whose language is unknown. However, little has been explored in terms of how electronic translators are used by readers, in addition to that, we do not know the strategies that are involved in this type of reading. Boruchovitch (2001), for example, investigates the repertoire of reading strategies and states that the lack of their use is a problem to be faced in reading. Therefore, we believe it is important to know the strategies when using machine translation and how they can help teachers as well as readers in the area of teaching/learning a L2 / FL.

In this study we are describing global and local strategies as well as cognitive and metacognitive strategies used by readers from different cultural backgrounds and languages, however, they all have the same goal which is reading in English. During reading comprehension the reader may translate unknown words, phrases that are difficult or the text altogether. While reading in the mother tongue is obviously a procedure done in the L1, when the translator is used reading becomes a hybrid process because both L1 and L2 are at stake.

Finally, reading comprehension is seen here as production of meanings from the text that is mostly accomplished through cognitive strategies. In the specific case of this study, when using the electronic translator reading comprehension must occur differently than when one reads without any support strategy. This means that the tool changes the way we process information.

1.2 Reading strategies

There is no common sense about the concept of reading strategies, and even less about their classification. Some authors distinguish global and local strategies, cognitive and metacognitive, others differ skills and strategies. More broadly, Weinstein and Mayer (1986) define cognitive strategies as actions that help govern behavior, emotion, motivation, communication, attention and understanding. However, this is a very general concept, and does not refer exclusively to reading strategies. Koda (2005) mentions the definition of Wasik and Turner- "actions deliberately selected to achieve particular goals" as that accepted by many researchers in the field.

A definition that seems quite clear is the one by Barnett (1989, p.66) cited below:

The term *strategy* refers to mental operations the reader purposefully approaches a text to make sense of what he\she reads. Strategies can be controlled by the reader's awareness or unconscious processes automatically applied.

According to this definition, the author sees strategies related to reading comprehension since the reader makes sense of what he\she reads. There are several works related to reading strategies in a L2, such as in Hosenfeld (1977), Anderson (1991), Li and Munby (1996), Andrade and Tomitch (2012), they all correlate strategies with reading comprehension. The study by Anderson (1991), for example, shows how participants who used more strategies had a higher comprehension.

The strategies analyzed in this investigation are based on the classification by Anderson (1991). In order to achieve the objectives of this research itself, Anderson's classification was adapted and a total of 26 reading strategies came up. Originally, Anderson(1991) included in his list a category of strategies used in tests, this is not the objective of this investigation, then, we excluded this category. Figure 1 below describes those strategies found in our own data:

Figure 1. List of reading strategies from the protocols. (Adapted from Anderson,1991)

<i>Supervision strategies</i>	<i>Support Strategy</i>
1 Recognizes loss of concentration	11 Skips unknown words
2 States failure to understand some part of the text	12 Needs a dictionary or translator
3 States success in understanding part of the text	13 Scans the material to search for a specific word
4 Adjusts reading speed to increase understanding	
5 Asks a question	

6 Makes a prediction about the meaning of a word or text content 7 Refers to a lexical item that prevents understanding 8 Confirms / disconfirms an inference 9 Refers to an earlier passage 10 Does self-repair <i>Paraphrase Strategies</i> 14 Uses cognates from L1 and L2 to understand 15 Breaks the lexical item in parts 16 paraphrases 17 Translates the word or phrase in L1 18 Summarizes	<i>Maintenance strategies of textual coherence</i> 19 Rereads 20 Uses clues of the context 21 Reads forward <i>Scheme-oriented Strategies</i> 22 Uses world knowledge 23 Recognizes lack of world knowledge 24 Relates the phrase to a personal experience 25 Extrapolates the information presented in the text 26 Guesses a word without any consideration
---	---

The figure above classifies strategies into five categories: Supervision, Support, Paraphrase, Coherence and Scheme oriented. We chose this classification once the five categories represent the processing of the strategy. The first type refers to how the reader monitors his or her reading, the second type deals with strategies used as an auxiliary aid to reading, as it is the case of the electronic translator. In the third category there are strategies for paraphrasing and the fourth type is related to maintenance of textual coherence. Finally, the last category of strategies are those that the reader makes use of world knowledge / schema for comprehending.

What researches in the field show is that the strategy itself is not better or worse for reading comprehension, but its efficient use is what makes it successful. In this regard, it is important to check how readers utilizing Google translator employ strategies. As for the translation strategy with Google's tool does the reader translate at the lexical level? the sentence? the text? It is not enough to know what the strategy is, the reader should be able to use it strategically. Moreover, as claimed by Huang, Chern and Lin (2008) it is more important to know how to evaluate the success of a particular strategy, as well as to learn how to correct it rather than just use it. We need to investigate how readers are using this support strategy, because what has been done so far is deny (by teachers) machine translation. That is also a consequence of the very lack of research about the subject.

1.3 Machine translation

Machine Translation is an area of research in natural language processing (PLN). This field of knowledge attempts to answer how machines can understand natural language so as to perform translation. As linguists, we know of numerous language problems faced by translators, machine translation is even more problematic, because it obviously does not think like humans. Some language issues (Mitkov, 2003) that cause problems are the disambiguation of the meaning of the word, anaphora and syntactic processing in general. On the one hand, research on reading strategies is a great potential for research, on the other hand research into new technologies for teaching/learning English as a L2/FL has also developed. There are many software for English

teaching, such as online courses where native and non-native speakers interact, online dictionaries, online linguistic corpus, and of course, automatic electronic translators. These technologies have helped and facilitated learning, however, the area of machine translation and its pedagogical impact has been the least researched. From this perspective, there is a need for more research which can contribute to the teaching area of foreign languages (with a psycholinguistic bias), especially the field of reading and the role of machine translation.

Not only in the university, but in society as a whole, there is a need to read and understand texts in English as a L2. However, the process of learning a foreign language up to the proficiency level is long. So, counting on the aid of an electronic tool that facilitates understanding is a benefit. It is logical to think that there are limits to comprehending translation made by machine, but it is necessary to investigate what these limits are and how effective the machine is. Historically, students use translation tools feeling there is comprehension, in this respect, what is the extent of this comprehension?

It is noteworthy that in this work we deal with the genre of academic text (abstracts), for this reason we see the possibility of using Google translator. The quality of this kind of translation is better than a more metaphorical or poetic text that would require more complex analysis the machine is not capable of doing. Yet, we recognize that there are problems with Google translation, but researchers are working to develop it. Many advances have been made in electronic corpora as a subfield of NLP (McEnery, 2003), as the area advances we are likely to see better translators.

The problems machine translation deals with range from syntax (e.g.. a more appropriate verb tense, the correct preposition...) to semantics or pragmatics (anaphoric decision, lexical ambiguity). Machine translation makes use of specific methodologies for addressing these issues. Unlike most commercial automatic translators that use rule-based systems, Google translator is based on the statistical approach to Machine Translation. Its corpus consists of all texts which were translated into several target languages published on the web. The advantage of this translator is that the texts were all translated by humans and this ensures more reliability in translation.

According to Lima (2011) from the three most popular electronic translators Google remains the most efficient. On Google's website the following information is available:

Our system takes a different approach: we feed the computer billions of words of text, both monolingual text in the target language, and aligned text consisting of examples of human translations between the languages. We then apply statistical learning techniques to build a translation model. We've achieved very good results in research evaluations.

In fact, a comparison among other on line translators such as *Bing* or *Babel fish* (Lima,2011) shows that Google offers better translation results. It is true that cultural, ideological and idiosyncratic issues in a language are difficult to be translated by Google, but they are investigations linguistics as a whole does not have a final answer.

METHOD

Procedures

Our aim here is to analyze strategies used by readers who translate either using Google translator or not. We carried out a qualitative research using the methodology *Think aloud*

protocols with 10 intermediate students at the ELI (English Language Institute) from the University of Pittsburgh. *Think aloud protocol* is a technique in which the reader expresses his/her understanding at the exact moment of reading. According to Caldwell (2008) this methodology has been widely used for providing valuable data on the cognitive strategies employed in an attempt to understand a text. It is suitable for the analysis of reading comprehension because the student through the explicit protocol verbalizes his/her reasoning.

The reason we chose students at intermediate level of proficiency was because advanced proficient students would already be able to read without a support for understanding. Furthermore, the methodology of Verbal Protocols would not assist the basic levels. These participants are all college students in Pittsburgh or in a university in their home countries, but they are all students of English as a L2 at the ELI.

The nationalities of the participants varied and four out of ten are PhD students. It is believed that the strategy of electronic translation should be a benefit for all languages available in Google translator. The reading instruments used referred to an abstract of a scientific article published in a journal on the Internet whose topic was about international politics. Since these students come from many countries to study in the United States we believed this topic would be of interest to all of them. Participants should read the abstract on the computer screen for two times and whenever there was a signal to pause they should explain their comprehension. The first time they could not use any support strategy and the second one they could use Google translator. Prior to the running of the instrument a questionnaire was conducted to verify the participants' reading habits. The intent was to check if student's reading habits have an impact in their comprehension (Perfetti:2007). Figure 2 below summarizes the profile of the participants.

Figure 2. Profile of the participants

Participants	10
Level of English	Intermediate
Education	4 PhD graduates 5 undergrads and 1 specialist
Translated languages	Chinese (1) Japanese (2) Arabic (4), Portuguese (2) Spanish (1)

There was no need to conduct tests based on text content questions because the technique of *Think aloud* is a methodology that provides data to analyze comprehension. The verbal protocols were recorded with the audio software *Audacity*. However, the student was not told whether to translate words, a sentence or the whole text. The purpose was to analyze the needs and preferences when using Google translate.

Participants read the same abstract for two times since the goal was to verify if Google helped the student with what he did not understand or did not know during the first reading. It would be necessary to identify the exact moment where there was no comprehension, so that we would know if Google was a benefit. The comparison between understanding / not understanding would only be possible with the two identical texts. The following figure (figure 3) is a summary of the instruments conducted.

Figure 3.Summary of instruments

Step 1 Questionnaire on reading habits (5')
Step 2 Think aloud protocol - (up to 15 ') Instrument 2: Reading and speaking about the text: <i>Is the European Commission too powerful? Neofunctionalism and intergovernmentalism considered.</i> Without using Google translator
Step 3 Think aloud protocol - (up to 15 ') Reading and speaking about the text: Is the European Commission too powerful? Neofunctionalism and intergovernmentalism considered. Using Google translator.

1

RESULTS AND DISCUSSION

We identified twenty-five and twenty six strategies in the analysis of the protocols (instrument 2 and 3). These results show that there is hardly any difference among strategies used when the participant read with or without the electronic translator. The difference is how participants managed these strategies. When they used Google strategies were used more often, that is, although the strategies are the same they were more frequent. This demonstrates that participants engage more strategies when using the translator, there are more cognitive demands. Furthermore, in general terms data demonstrates that comprehension was more effective when the participants used the tool.

In order to facilitate the analysis of the results see the summary of data found in the figure below:

Figure 4. Data

	INSTRUMENT 2	INSTRUMENT 3
Total of strategies	25	26
Most used category	Support	Supervision, support and paraphrase
Most used strategy	Skips unknown words	Refers to lexical items that prevents understanding, requires use of the translator, translates, expresses success

Most used strategy by PhD students	Skips unknown words and summarizes	of understanding skips words and summarizes refers to the lexical item that prevents comprehension, needs to use the translator, translates, expresses success of understanding
Most used strategy by those who prefer Google	Confirms or disconfirms an inference	Confirms or disconfirms an inference
Most used strategy by speakers of western languages	Skips unknown words and summarizes	skips words and summarizes, Refers to the lexical items that prevents understanding, requires the use of translator, translates, expresses understanding of success
Most used strategy by speakers of eastern languages	Skips unknown words	refers to the lexical item that prevents understanding, requires use of the translator, translates, expresses understanding of success, rereads

Regarding the types of strategies that were mostly used in the instrument with Google translator the ones related to supervision and paraphrasing appear. This suggests that the use of the digital tool make participants to vary more and be less repetitive. It has been said previously in this article that it is the combination of strategies that lead to success in comprehension. Thus, the range of options suggests that the participant is not restricted to the same mental operations and this ensures more understanding. Supervision, for example, monitors reading, so the participant is attentive to his/her own performance. Thus, machine translators seem to make readers aware of their understanding as well as make them realize when there is no comprehension.

Concerning strategies used by PhD candidates results of the two instruments differ greatly. The first time participants read the text without Google they skipped the unknown words and paraphrased more, whereas when using the translator they developed a pattern that is referring to a lexical item that prevented comprehension which resulted in the use of the translator. After translating, participants expressed success in reading comprehension. Accordingly, when reading with Google they used bottom up processing in order to achieve local understanding and then top-

down processing which enabled to understand the sentence or the passage. This suggests that comprehension occurs in a less superficial level rather than only skipping unfamiliar words or paraphrasing. Therefore, the tool changes the way participants process information.

According to the results described above, most PhD students in instrument 2 paraphrased more as well as simply skipped unknown words. This kind of reading is related to a more superficial level that refers to extracting main ideas but not the subtleties of the text. As for the other instrument these participants refer to the lexical item that prevents understanding, so they express the need to use the translator. When translation is provided the reader can insert the correct meaning to the sentence or text. What we mean is that it is not only translation the cause of comprehension, but recognizing there is lack of comprehension and then knowing how to use translation in the context of the text. This combination of strategies would lead to an interactive processing both top down and bottom up, which ensures greater comprehension.

With reference to the strategy mostly used by those who prefer to read with the electronic translator it does not differ from one instrument to another. In both contexts participants confirmed or disconfirmed their inference. The difference is that with Google translator they could confirm their inference more than disconfirm. Thus, we believe there was more understanding in instrument 3. Literature in the area says that when the reader infers there is a deeper comprehension. Also, as we mentioned previously, inferential abilities are at a higher level of comprehension. This result showed us that the ones who used Google were better comprehenders than the others. We point out that even when the participant disconfirms the inference his/her metacognition gets into action because he/she becomes aware of it. If the participant simply skips the words he/she identifies as a problem, reading comprehension is likely to be less effective and superficial. Likewise, if there are many words which are simply ignored understanding must be very poor.

Finally, strategies used by participants of Eastern and Western languages in instrument 2 were similar. However, in instrument 3 speakers of Chinese, Japanese and Arabic also reread after all. This rereading shows that they might want to ensure understanding or they were not completely sure if they really comprehended. We have to take into consideration that these Eastern languages are distant from the Indo-European, so there are more cognitive demands from the reader. We cannot say there is less comprehension, but we can say they put more effort to perform the reading task.

DISCUSSION AND CONCLUSION

In general, the analysis of reading strategies made us believe that Google translation was a benefit to reading comprehension, and that such strategies are managed differently when using this support. As for grammatical problems of machine translation it is questioned if they are an impediment to comprehension as it has always been believed. Even if many syntactical problems occur, such as the ones seen in the protocols of the Arabs, they can still edit the text to comprehend. Grace (1998) says the reader wants equivalence between L1 and L2.

In this research participants who used Google could monitor their reading comprehension more, it seems they used metacognitive strategies more frequently. Unlike the dictionary, the translator provides access to the meaning of the words more rapidly too. Because of this access there seems to be more automaticity to read, which frees memory to perform other activities. Yet, the reader has to scan the word in the dictionary which is another cognitive demand. Scanning is not necessary with the translator. Thus, when the translator is used working memory is less loaded, because there is less interruption of unknown words or guessing from the context.

Data shows that meanings were constructed much from the words participants translated. For example, the keyword *comission* assured comprehension of the general idea of the text. Moreover, the words *neofunctionalism* and *intergovernmentalism* could be comprehended much more from the clue of *schools of thought*.

We have to take into consideration that participants were leveled intermediate, and had not reached a final attainment yet. So, much of the extent of comprehension counted on Google´s machine.

Finally, it is worth explaining that the reader / learner should not be dependent on the translator, it should be used as a support tool for students. Importantly, Google translator is not a threat to language teachers, rather, its role is to help teachers to address where translation makes sense or not. As described by Caldwell (2008) understanding and learning are distinct concepts. Learning involves remembering and applying components that cannot be part of the comprehension process, we can understand and then forget. However, learning involves long-term memory and high level of awareness and attention. We cannot guarantee the participant has learned the words translated by Google (e.g. *comission*, *comitology*, *neofunctionalism* and *pillar*). This means that Google itself does not teach. However, its use enhanced understanding, which generated new knowledge about the topics. Nevertheless, the teacher can highlight important lexicon to be remembered, core vocabulary as well as work on grammar topics which are relevant.

One of the most important conclusions that can be reached is that using Google translator decreases the difference between readers who know more and those who know less. Although the participants had been leveled intermediate by the ELI, individual differences were perceived during the protocols. For example, two participants had more problems to comprehend the text. Also, the one who refused to use the tool did not show any improvement when reading for the second time. Yet, we realized comprehension was reached at different levels by the different readers.

The contribution of this research points out to the use of Google translator as a support strategy which can be beneficial when other strategies are at stake. Altogether they can promote and accelerate comprehension by readers at different levels of knowledge. Thus, both intermediate and basic level students could master an advanced student comprehension if they read strategically.

REFERENCCEES

- Anderson, N. J. (1991). Individual differences in strategy use in second language reading and testing. *The modern language journal*, 75(4), 460-472.
- Andrade,A. Gil,G. Tomitch,L.(2012) Percepção de estratégias de leitura em LE de alunos universitários. *Revista (Con)Textos Linguísticos • Vitória – v. 6, n. 6 p. 7 – 17.*
- Barnett, M. A. (1989). *More Than Meets The Eye: Foreign Language Reading. Language and Education: Theory and Practice.* Prentice-Hall Regents, Englewood Cliffs, NJ 07632..
- Boruchovitch, E. (2001). Algumas estratégias de compreensão em leitura de alunos do ensino fundamental. *Psicologia Escolar e Educacional*, 5(1), 19-25.
- Caldwell, J. S. (2002). *Reading Assessment: A Primer for Teachers and Tutors.* Guilford Publications, Inc., Dept. 3R, 72 Spring St., New York, NY 10012.
- Grace, C. A. (1998). Retention of Word Meanings Inferred from Context and Sentence-Level Translations: Implications for the Design of Beginning-Level CALL Software. *The Modern Language Journal*, 82(4), 533-544.
- Hosenfeld, C. (1977). A preliminary investigation of the reading strategies of successful and nonsuccessful second language learners. *System*, 5(2), 110-123.
- Huang, H.C.,Chern, C.L., & Lin, C. (2008).EFL learners' online reading strategies: A comparison between high and low EFL proficient readers. *English Teaching and Learning*, Special Issue (1), 1-22.
- Kintsch, E. (2005). Comprehension theory as a guide for the design of thoughtful questions. *Topics in Language Disorders*,25(1), 51-64.
- Koda, K. (2005). *Insights into second language reading: A cross-linguistic approach.* Cambridge University Press.
- Li, S., & Munby, H. (1996). Metacognitive strategies in second language academic reading: A qualitative investigation. *English for specific purposes*,15(3), 199-21
- Lima, L. (2011). Tradução eletrônica: do riso irônico ao interesse científico. *Fólio Revista de Letras*,3(1).
- McEnery, A. (2003). Corpus linguistics. In: R.Mitkov (ed.) *The Handbook of Computational linguistics.*Oxford University Press.
- Mitkov, R. (Ed.). (2003). *The Oxford handbook of computational linguistics.* Oxford University Press.
- Perfetti, C. (2007). Reading ability: Lexical quality to comprehension. *Scientific studies of reading*, 11(4), 357-383.
- Weinstein, C. E., & Mayer, R. E. (1986). The teaching of learning strategies. *Handbook of research on teaching*, 3, 315-327.