



An empirical investigation of cognitive effort required to post-edit machine translated metaphors compared to the translation of metaphors

Arlene Koglin

LETRA, Universidade Federal de Minas Gerais

arlenekoglin@yahoo.com.br

DOI: [ti.106201.2015.a06](https://doi.org/10.106201.2015.a06)

Abstract: This paper reports on a study that investigates the cognitive effort required of a translator either to post-edit machine translated metaphors or to translate metaphors manually by means of eye tracking and keystroke logging. We hypothesise that post-editing will be less effortful than manual translation. In order to test this hypothesis, an experiment was conducted with two different groups of participants. Fourteen participants were asked to post-edit a newspaper text and eight participants were asked to translate the same source text. The analysis focuses on eye-tracking data related to total fixation duration and keystroke logging data (insertions, deletions, pauses). Data analysis shows that the cognitive effort required to post-edit MT output is lower in comparison to manual translation.

Keywords: cognitive effort, metaphor, translation, post-editing, eye tracking

1. Introduction

Machine translation has been developing rapidly in recent years; however, we cannot disregard the role of post-editors in order to correct the errors in the machine translation output. This task is commonly referred to as post-editing. According to Allen (2003, p. 296), the activity of post-editing is characterised by “editing, modifying and/or correcting a pre-translated text that has been processed by a machine translation system from a source language into a target language.”

The machine output quality may be affected by factors such as the text type to be translated, the machine translation system that has been implemented, or if controlled language has been used during authoring. Both MT quality and the types of errors produced can be the result of either source text features or related to the machine translation system itself. For example, if the MT system is statistically based, then semantic content may be machine translated more successfully, whereas a rule-based MT system will often be grammatically superior.

Furthermore, the text type may pose additional difficulties for the machine translation if it is characterised by long, complex sentences or metaphorical utterances. As a result, the poorer the machine output quality, the more post-editing effort will be needed — provided that the text is expected to be fully post-edited.

Over the past decade most research and post-editing guidelines have encouraged the use of post-editing with technical texts, which are likely to have greater terminological consistency and simpler sentences. Recent evidence on post-editing (Carl et al., 2011) suggests that the machine

translation of newspaper texts with subsequent post-editing is faster than manual translation.

The findings of Carl et al. (2011) may be encouraging with regard to machine translating newspaper texts, but additional investigation is needed into the effort required to post-edit them. One of the reasons that findings regarding post-editing effort of technical texts might not apply to post-editing newspaper texts is because the latter are characterised by the use of metaphorical language, which may require additional effort during the post-editing task.

A considerable amount of literature has been published on metaphor processing and effort required to interpret metaphors, but there is no firm evidence of its role on post-editing effort. Yet it is becoming increasingly difficult to ignore this issue due to the improvements of machine translation and the growth of the post-editing market.

This paper seeks to address the question of cognitive effort involved in the post-editing of metaphors. To do so, the study compares the cognitive effort required to post-edit machine translated metaphors and to translate metaphors. We hypothesise that manual translation will require more cognitive effort than post-editing.

2. Theoretical framework

2.1 Metaphor processing effort: a relevance-theoretic approach

Relevance theory claims that every act of ostensive communication is guided by the presumption of relevance, which enables people to draw inferences from the given stimulus. As our “mind tends to allocate its resources to the most relevant information, if the communicator wants to be understood, he should produce a stimulus which is at least relevant enough to the interpreter to be worth” processing (Unger, 2001, p. 29). The interpreter can, in turn, “interpret the stimulus on the assumption that it will be at least adequately relevant to him. This justifies acceptance of the first accessible interpretation which satisfies his expectations of optimal relevance.” (Unger, 2001, p. 29)

Consequently, relevance theorists see metaphor differently than cognitive linguists, insofar as metaphors in this theoretical approach are a matter of language use. Moreover, they claim that metaphor is not unique as its use is just another form of loose interpretation (Sperber & Wilson, 2008). Under the relevance-theoretic view, metaphors do not necessarily require more effort to be processed given that they are interpreted¹ through pragmatic processes of loosening and narrowing. Instead, relevance theorists maintain that metaphorical language “reveals important insights into the cognitive and communicative principles that motivate people’s striving for optimal relevance in interpersonal situations.” (Gibbs & Tendahl, 2011, p. 602)

Since the mind tends to allocate its resources to the most relevant information, the mind should produce a stimulus which is at least relevant enough to the addressee to be worth attending to, provided that the communicator wants to be understood. The addressee can therefore interpret the stimulus on the assumption that it will be at least adequately relevant to him. This justifies acceptance of the first accessible interpretation which satisfies his expectations of optimal relevance.

According to Gutt (1992), the inferential approach of relevance theory allows us to have a more precise understanding of translation. Based on the notion of interpretive resemblance, i.e., the interpretation of meaning relies on

¹ The terms “process” and “interpret” are used interchangeably in this paper in order to be consistent with terminology used in relevance theory.

resemblance in semantic representation or logical form (Sperber & Wilson, 1986, p. 228), Gutt (1992) addresses translation as the re-creation of logical and inferential properties of source text in the target text.

Under the relevance-theoretic account, the human cognitive system aims at maximising relevance; therefore the interpretation of any ostensive stimulus is expected to demand the investment of the smallest possible cognitive effort to derive the greatest possible contextual effects (meanings) (Sperber & Wilson, 1986).

This balance between maximising contextual effects and minimising cognitive effort is a procedure of interpretation claimed by relevance theorists to explain both metaphorical and non-metaphorical interpretation equally. That is, metaphor interpretation constitutes a pragmatic phenomenon in which the interpreter will stop the process when the interpretation satisfies the principle of optimal relevance.

However, there is no agreement among scholars regarding the cognitive effort required to interpret metaphors. The challenging question is: Do metaphors take longer to be processed in comparison to non-metaphorical utterances?

One clue to answer this question comes from psycholinguistic experimental findings which showed that people took “significantly longer to read the metaphors in the contradictory contexts (1939 milliseconds) than they did either the strengthening (1717 milliseconds) or contextual implications (1709 millisecond) contexts” (Gibbs, Tendahl, & Okonski, 2011). These findings suggest that context and not the type of metaphor (creative or conventional) determines the amount of effort allocated to metaphor processing (Gibbs & Tendahl, 2008).

Under the relevance-theoretic account, context is understood as “the set of premises used in interpreting an utterance” (Sperber & Wilson, 1986, p. 15) — therefore it comprises a psychological notion. *Context* is a term used with different meaning in comparison to other theoretical notions. Sperber and Wilson clarify the notion of context as follows:

A context in this sense is not limited to information about the immediate physical environment or the immediately preceding utterances: expectations about the future, scientific hypothesis or religious beliefs, anecdotal memories, general cultural assumptions, beliefs about the mental state of the speaker, may all play a role in interpretation. (Sperber & Wilson, 1986, p. 15–16)

More precisely, the notion of context refers to a part of the cognitive environment of the addressee. The individual’s cognitive environment, in turn, consists of all the facts that he is capable of representing in his mind. The sources of this information can be perception, memory, or inference, which can be made based on the two previous sources (Gutt, 1992).

According to Gibbs (2010), metaphor interpretation is not a singular and simple process, but rather depends on the person’s goals or the task in which he is engaged. Nevertheless, there are few empirical studies that have investigated how the speaker’s pragmatic purpose influences both the process of metaphor understanding and the social meanings inferred by listeners (Gibbs, Tendahl & Okonsky, 2011).

The previous discussion justifies the importance of investigating the cognitive effort required to post-edit and to translate metaphors, because both activities are socially situated and have clear pragmatic purposes.

2.2 Post-editing effort

One of the main reasons for using post-editing in the market is the desire to save time and increase productivity. The chances of reducing time to post-edit and increasing productivity are closely related to the machine output quality.

Some discursive and textual aspects of the source text, such as ambiguities, long complex sentences, anaphoric referents, and metaphorical meanings represent additional difficulties for the machine translation system. Consequently, post-editing has been mainly used with technical texts which are characterised by less frequent occurrences of metaphors and less variation in terminology. Technical texts typically have less variation in terminology because controlled authoring and terminology management help to mitigate terminological inconsistencies.

Due to the development of machine translation systems, we believe it would be worth investigating whether newspaper texts are suited for post-editing because they are characterised by a relatively high percentage (approximately 15%) of metaphors (cf. Steen et al., 2010). One possible way of doing that would be examining the effort required to post-edit newspaper texts in comparison to manually translating them.

Krings (2001), who was a pioneer in investigating post-editing effort, classifies three different, but related, categories of post-editing effort as the key elements to determine if post-editing machine translation is worthwhile. The three categories are each described in turn.

2.2.1 Temporal effort

This concept is the easiest to measure because it refers to the amount of time needed to post-edit the machine output. If the post-editor saves time compared to human translators then using post-editing may be recommended.

2.2.2 Technical effort

Technical effort refers to the actual linguistic changes to correct the machine translation errors. In other words, it involves the process of deleting, reordering, inserting or a combination of all these actions to correct the errors. For example, the more insertions and deletions required to correct a text, the more technical effort to post-edit it.

2.2.3 Cognitive effort

This concept comprises the “type and extent of those cognitive processes that must be activated in order to remedy a given deficiency in a machine translation” (Krings, 2001, p. 179). According to Krings (2001), the cognitive effort is directly related to the previous concepts. Among the three concepts, this is the most complex to measure because it requires especial tools such as Translog or eye trackers, which do not measure cognitive effort directly, but are assumed to provide measures that represent it.

For more reliable results, Alves (2003) recommends data triangulation, which within the context of translation process research, could be the combination of eye-tracking and key-logged data and retrospective think-aloud protocols. In the same vein, O’Brien (2007) encourages researchers to triangulate technical and temporal effort analysis with pause analysis in order to have more reliable results and a deeper understanding of post-editing effort.

2.3 Pauses

In addition to Krings’ proposal of using the measures of temporal, technical and cognitive effort, pauses could also be used as an indicator of cognitive effort. According to Schilperoord (1996), analysing pauses during text production provides insights about cognitive processes and the focus of cognitive attention. Besides providing evidence of cognitive effort, pauses can

be a parameter for measuring the feasibility of both machine translatability of the source text and post-editing effort (O'Brien, 2006b).

However, when analysed in isolation, O'Brien (2006b) found that pauses are not reliable indicators of post-editing effort. Therefore, to have a deeper understanding, she suggests triangulating pause analysis with technical and temporal effort analysis. Besides that, O'Brien (2006b) concluded that more studies were needed in this domain in order to clarify how useful pause analysis is for both translation and post-editing research.

After discussing the concepts of temporal, technical, and cognitive effort related to post-editing effort, this section now moves on to consider the cognitive effort required to post-edit machine translated metaphors. Despite the considerable amount of studies that have been published on metaphor processing and effort required to interpret it, there is no firm evidence of its impact on post-editing effort.

3. Data collection

To address the gap in the literature regarding the cognitive effort required to post-edit texts rich with metaphors, we conducted this study to better understand the post-editing task in comparison to manual translation. We expected manual translation to be more cognitively demanding than post-editing.

3.1. Post-editing participants

The participants in the post-editing experiment were 14 undergraduate students at Federal University of Minas Gerais (UFMG) who volunteered to take part in the experiment. They all are native speakers of Brazilian Portuguese and consider English as their second language. Participants self-reported this information on a survey they were asked to answer before the experiment. They had no professional experience with post-editing but they all attended a 15-week post-editing course held at UFMG as part of the regular undergraduate course.

3.2 Manual translation participants

The participants in the manual translation experiment were 7 translators who also volunteered to participate in the experiment. These participants also were native speakers of Brazilian Portuguese and had English as their second language. Each participant had professional translation experience ranging from five to ten years.

3.3 Experimental design

Building on the experimental paradigm of data triangulation in translation process research, an experiment was conducted at the Laboratory for Experimentation in Translation (LETRA) using eye tracking, keystroke logging, and retrospective think-aloud protocols. First, all of the participants were asked to complete a short typing task in order to get familiar with all the keys on the keyboard. Next, seven out of the fourteen participants were asked to post-edit a target text that was machine translated using Google Translate in Task 1 (T1) and to post-edit a Systran machine translated output in Task 2 (T2). The other seven participants were asked to post-edit the same source text in a different order, i.e., Systran machine translated output in Task 1 (T1) and Google Translate in Task 2 (T2).

Seven different participants were asked to translate from scratch the same source text. At the end of each task, participants were asked to record the think-aloud protocols. In the first protocol, they were told to think aloud while

their full post-editing process was replayed on Translog-II screen. In the guided protocol, they were asked two questions related to metaphor interpretation and its post-editing decision-making process.

3.4 Material

Both tasks were performed using the same source text, i.e., a 224-word journalistic text about the Tea Party Movement (see Appendix A).

3.5 Procedure

In the post-editing experiment, participants identified with odd numbers were systematically assigned to receive Google Translate output for task 1 and Systran output for task 2. Participants identified with even numbers, on the other hand, had the opposite order of stimuli.

3.6 Apparatus

The participants were seated in front of a Tobii T60 eye tracker at a distance of 55 to 65 cm from the monitor. Both Translog-II and Tobii Studio 3.2 were calibrated. Translog-II enabled participants to view the source text in the upper half part of the window and the machine output in the lower half part of the window. This is software especially designed for process-driven studies because it enables tracking keyboard activity and mouse clicks.

4. Data analysis

For the purposes of this paper, the analysis of cognitive effort will focus on both key-logged data regarding pauses and eye-tracking data related to total fixation duration in two areas of interest (AOIs): in the source text (AOI1) and the target (AOI2) text. Both areas contained metaphors, The Tea Party Pork Binge and pork-barrel spending (cf. Steen et al. 2010).

Temporal effort was measured by the total time spent by each participant to complete the task, whereas technical effort was measured by the number of insertions and deletions (text production).

Due to poor quality eye-tracking data, two participants (P06 and P08) from the post-editing experiment and one (P01) from the translation experiment were discarded for the purposes of this analysis. The threshold set for eye-tracking data quality was 70% of time spent looking at the eye-tracker screen (cf. O'Brien, 2009). It is important to note, however, that no changes have been made to the participant names for two reasons: a) the number provides information on the machine translation output (see section 3.5 for more details) and b) both data and metadata from these two experiments will be freely available at the CRITT Translation Process Research (TPR) Database for future analysis.

All statistical analysis was performed using SPSS statistical software. The cut-off point for significance level was set at 0.05.

5. Results and discussion

5.1 Temporal effort

The participants from the current study did not have time constraints to perform the tasks of post-editing and manual translation.

Table 1 shows the distribution of the time spent on manually translating the newspaper text compared to post-editing it.

Table 1: Time spent on post-editing and translation tasks

Post-editing		Manual translation	
Participants	Duration (h:min:s)	Participants	Duration (h:min:s)
P01	0:25:16	P02	1:29:38
P02	0:12:57	P03	0:37:33
P03	0:30:51	P04	2:50:00
P04	0:12:42	P05	1:16:39
P05	0:51:23	P06	1:45:21
P07	0:40:48	P07	1:34:44
P09	0:43:37		
P10	0:58:02		
P11	0:45:23		
P12	1:09:21		
P13	1:45:32		
P14	1:17:29		

As can be seen in Table 1, the two distributions of time spent on each task are quite different. On average the text was post-edited in 47 minutes and 47 seconds while the manual translation took 1 hour 35 minutes 39 seconds.

The data also show that there is great variation among subjects. P04, for example, was the fastest post-editor followed by P02. On the other hand, the speed for P13 was above the mean time spent by translators. These results therefore need to be interpreted with caution. If the average time is taken into consideration, then post-editing of newspaper texts seems to save time in comparison to manually translating them. The Mann-Whitney U Test showed that time spent on post-editing was significantly lower ($Z = -2.154$, $p = .03$) than for manual translation. These findings are consistent with those of other studies (Krings, 2001; O'Brien, 2006a, 2007; Carl et al., 2011; Green, Heer & Manning, 2013) and suggest that post-editing newspaper texts may save time.

5.2 Technical effort

Technical effort is gauged here by measuring the number of deletions and insertions performed by each participant in both tasks: post-editing and manual translation. These measures are automatically provided by Translog-II statistics. It is important to note, however, that if the participant highlights a word and then deletes it, Translog-II will count this action as one deletion in the log file (xml).

Table 2 provides an overview of the number of deletions and insertions for each participant in the two tasks.

From the table we can see that there is considerable variation among participants regarding deletions and insertions. The mean value for deletions is lower for post-editing ($M = 193.17$, $SD = 96.35$) in comparison to manual translation ($M = 230.83$, $SD = 73.96$). Similarly, the average number of insertions, i.e., text production is lower for post-editing ($M = 620.17$, $SD = 380.45$) when compared to manual translation ($M = 1985.67$, $SD = 204.27$).

Regarding insertions, there was a significant difference ($Z = -3.372$, $p = .001$) between the two groups. This result can be obviously explained by the fact that translators perform the translation from scratch whereas post-editors only correct the errors from the machine output.

We would have expected that post-editors produce more deletions than translators, since they make changes on a machine translation output. Surprisingly, there is no significant difference ($Z = -.843$, $p = .44$) between deletions during post-editing and manual translation.

Table 2: Total number of insertions and deletions made by each participant on post-editing and translation tasks

Post-editing			Manual translation		
Participants	Deletions	Insertions	Participants	Deletions	Insertions
P01	135	847	P02	219	1872
P02	47	230	P03	165	1710
P03	154	319	P04	261	2158
P04	124	194	P05	130	1836
P05	265	581	P06	325	2183
P07	200	264	P07	285	2155
P09	126	680			
P10	329	779			
P11	88	844			
P12	337	1548			
P13	296	429			
P14	217	727			

There are some other possible explanations for the previous result. It might be that the post-editors judged the machine translation output as of good quality and consequently made fewer changes. Another possible explanation is that the participants may have followed two important post-editing guidelines: retain as much raw translation as possible and make changes only where absolutely necessary, which were part of the instructions given before the task and during the course that the participants took part.²

Besides that, the higher number of deletions made by the group of translators may be a result of typing errors. It might be possible that these participants were not very skilled at typing and had to delete a great amount of typing errors. In order to test this hypothesis, a detailed analysis of the typing activity per minute performed during the typing task was conducted. The typing activity was correlated with the number of deletions performed by the group of translators during the translation task.

A Spearman's correlation coefficient was run to determine the relationship between the two variables, but the result failed to reach statistical significance ($r_{sp} = -.771, p = .07$). The p value suggests, however, a significant trend towards a negative correlation between the number of deletions and the translators' typing skills, i.e., the better the participants are at typing, the fewer deletions made during the task. Despite the small sample size ($n = 6$) and the marginally significant result, we may suggest that a negative correlation between the number of deletions and the typing activity could have implications for the validity of this measure as an indicator of technical effort.

5.3 Cognitive effort

In order to compare cognitive effort required to post-edit and to translate from scratch, fixation duration, i.e., the time that a fixation lasts, on two metaphors was calculated. Longer fixations represent higher cognitive effort.

Figure 1 shows fixation duration in seconds while participants were post-editing and translating The Party Pork Binge (metaphor 1 – henceforth M1).

² The general post-editing guidelines (cf. O'Brien, 2009) provided for the students were as follows: a) Retain as much raw translation as possible, b) Do not hesitate too long over a problem, c) Do not worry if style is repetitive, d) Do not embark on time-consuming research and e) Make changes only where absolutely necessary.

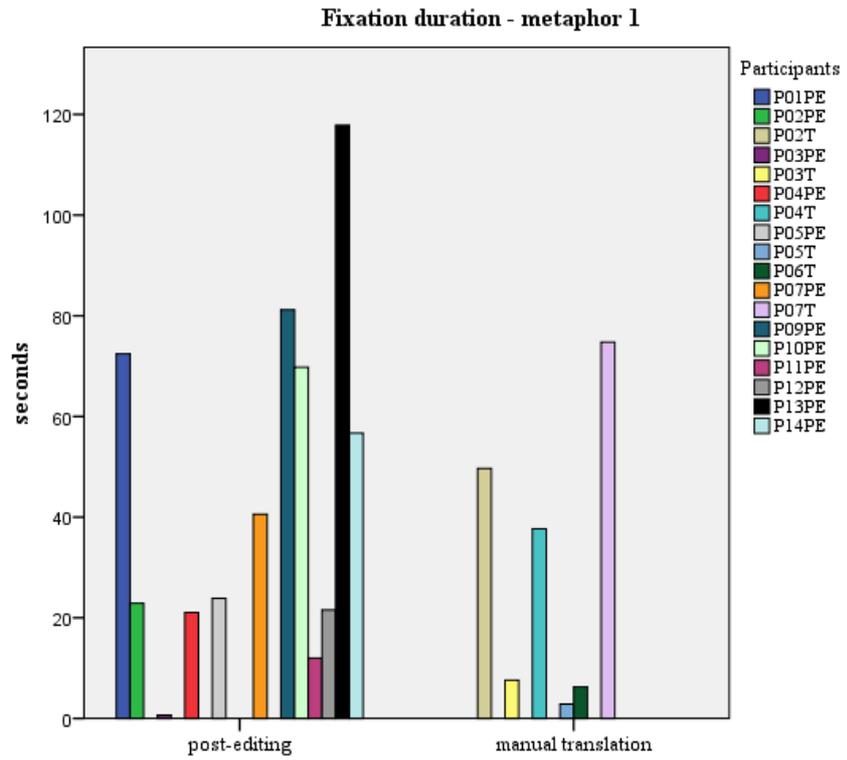


Figure 1: Comparison of total fixation duration in seconds in the post-editing and manual translation of metaphor 1 (M1).

Figure 2 shows fixation duration in seconds while participants were post-editing and translating the metaphor pork-barrel spending (metaphor 2 – henceforth M2).

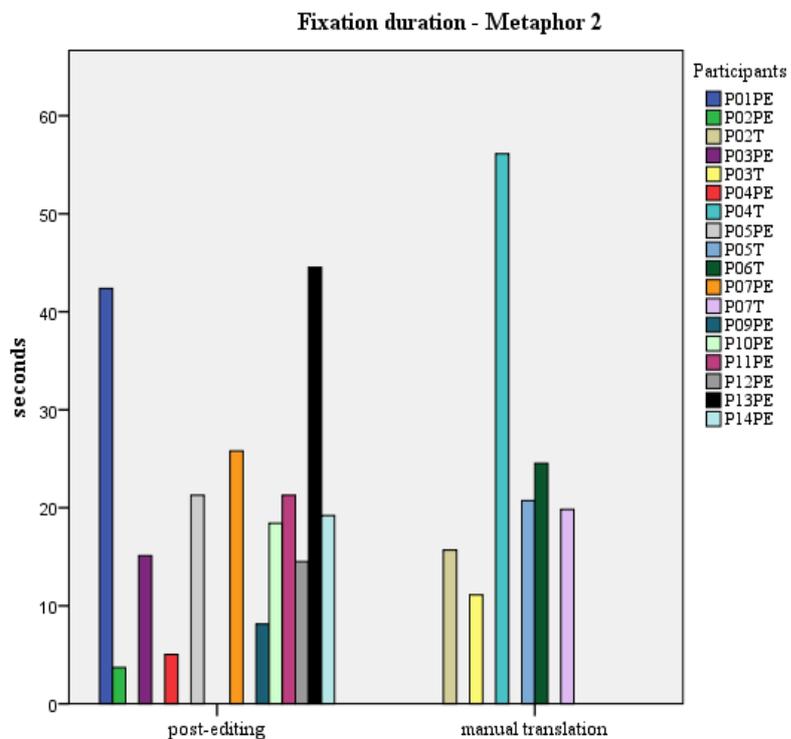


Figure 2: Comparison of total fixation duration in seconds in the post-editing and manual translation of metaphor 2 (M2).

As can be seen in the figures, the comparison of post-editing and manual translation behaviour with respect to fixation duration shows reasonable variation among participants of both tasks. From a theoretical standpoint, this seems to be compatible with the idea that metaphor interpretation may be influenced by participant's beliefs and motivations (Gibbs & Tendahl, 2008).

Alternatively, variation among participants could be a result of either their different cognitive environments (Alves, 2005) or their goals during their engagement in both tasks (Gibbs, 2010).

In order to compare the average fixation duration in post-editing, Table 1 shows the means for M1 and M2.

Table 3: Mean fixation duration in seconds at M1 and M2 of post-editing and manual translation

	Post-editing		Manual Translation	
	AOI1	AOI2	AOI1	AOI2
M1	11.02 (<i>SD</i> = 13.85)	34.02 (<i>SD</i> = 24.14)	15.46 (<i>SD</i> = 12.74)	14.34 (<i>SD</i> = 18.43)
M2	5.51 (<i>SD</i> = 5.41)	14.44 (<i>SD</i> = 8.50)	14.23 (<i>SD</i> = 14.11)	10.44 (<i>SD</i> = 2.83)

Differences between fixation duration in terms of source text (ST) and target text (TT) distributions show that participants had longer average fixation durations on the TT area in post-editing and longer average fixation durations on the ST area in manual translation. These results suggest differences in processing information at those tasks.

Post-editing results regarding distributions in terms of ST and TT are similar to those found by Carl et al. (2011), i.e., participants fixated longer in the source text when manually translating it, whereas participants who did post-editing fixated longer in the target text. The authors (2011, p. 140) explain that post-editing process encompasses “first reading a segment of raw SMT output, then comparing this against a segment in the ST that it is a translation of, and then possibly correcting the machine- translated output and reading the corrected version one or several times.” We have to note that all our post-editors attended a 15-week practical and theoretical post-editing course, whereas none of the participants from the previous study had experience in post-editing.

In manual translation, participants fixated for longer periods of time in the ST than in the TT, which indicates more cognitive effort at ST. This is probably due to the need of a deeper understanding of the ST. Our results are different from those found by Carl et al., in which TT had longer fixation duration. The reason for this finding is not clear but it might be related to differences in the experimental design of the two studies. In the experiment carried out by Carl et al. (2011) the participants had to translate under time constraints, therefore they may have avoided re-reading the source text and fixating longer as a time-saving strategy.

In relation to M1 and M2, Table 3 shows that fixation duration on the ST area (AOI1) was longer in the manual translation. It may be that the participants benefited from the machine translation output to infer the metaphorical meanings.

Contrary to expectations, the average fixation duration on the TT area (AOI2) for both M1 and M2 was longer in the post-editing. However, the Mann-Whitney U Test showed that these results were not significant either in M1 ($Z = -1.780, p = .07$) or in M2 ($Z = -1.218, p = .22$). Together these results provide important insights to understand the process of post- editing in comparison to manual translation with a focus on texts rich with metaphorical meanings.

Let us now turn to the next section, which is concerned with the analysis of pauses during the process of post-editing and translation, since they are also indicators of cognitive effort.

5.4 Pauses

Pauses are indicators of post-editing effort as well as a key element to assess the usefulness of machine translation output (Klings, 2001; O'Brien, 2006b).

Table 4 provides an overview of the total of pause duration for each participant.

Table 4: Comparison of total of pause duration in the post-editing task.

Post-editing		Manual translation	
Participants	Pauses (min.)	Participants	Pauses (min.)
P01	16.27	P02	67.01
P02	8.08	P03	27.69
P03	26.33	P04	84.39
P04	10.16	P05	61.93
P05	45.05	P06	74.40
P07	35.51	P07	80.88
P09	38.72		
P10	46.74		
P11	38.79		
P12	56.14		
P13	80.56		
P14	68.34		

These data show that the total of pauses is lower in the post-editing task ($M = 39.21$, $SD = 22.27$) when compared to manual translation ($M = 66.05$, $SD = 20.57$). The Mann-Whitney U Test showed that the difference is significant ($Z = -2.154$, $p = .03$).

This is another favourable result regarding the post-editing of texts that are rich with metaphors. Besides, the pause analysis from the current study produced results which corroborate the findings of previous work in this field (Klings, 2001).

Further studies, which analyse pauses in metaphorical utterances in comparison to non-metaphorical utterances, will need to be undertaken in order to determine the amount of effort required to post-edit metaphors.

6. Conclusions

The present investigation has compared cognitive effort required to post-edit and to translate from scratch newspaper texts. The preliminary findings of this study suggest that post-editing could be less effortful than manual translation when the following variables: pauses, task duration and insertions are taken into account. These findings on post-editing effort corroborate the results from previous studies and contribute additional evidence that indeed post-editing saves time. However, with a reasonably small sample size in the control group and high variation among participants, caution must be applied, as the findings might not be generalisable to all translators.

Although further work will need to be done to investigate thoroughly metaphor post-editing effort, our findings would seem interesting in that they challenge long-established assumptions about the unlikelihood of post-editing texts rich with metaphors. There has not been a significant difference between

cognitive effort required to post-edit machine translated metaphors in comparison to manually translating them. This result may be due to the small sample size of the control group or the small number of metaphors analysed in this paper. Therefore, further analysis with a higher number of metaphors should be conducted in order to have a more thorough understanding regarding cognitive effort involved in post-editing and manual translation.

Acknowledgements

We would like to thank all the participants who kindly accepted to contribute to this research project. The author also would like to thank the National Council for Scientific and Technological Development (CNPq) for financial support.

About the author

Arlene Koglin is a PhD candidate in Translation Studies at the Federal University of Minas Gerais (Brazil). She holds a Master's degree in Translation Studies from the Federal University of Santa Catarina (Brazil). Her current research and publications focus on post-editing, translation process, metaphor, cognitive effort and eye tracking. She has experience teaching post-editing and translation to university students.

References

- Allen, J. (2003). Post-editing. In H. Somers (Ed.). *Computers and Translation: a translator's guide* (pp. 297–318). Amsterdam-Philadelphia: John Benjamins.
- Alves, F. (2003). Tradução, cognição e contextualidade: triangulando a interface processo produto no desempenho de tradutores novos. *Translation, cognition D.E.L.T.A.*, 19, 71–108.
- Carl, M., Dragsted, B., Elming, J., Hardt, D., & Jakobsen, A. L. (2011). The Process of Post-Editing: a pilot study. In B. Sharp, M. Zock, M. Carl, & A. L. Jakobsen (Eds.), *Proceedings of the 8th International NLPCS Workshop: Human-Machine interaction in translation* (pp. 131–142). Copenhagen: Samfundslitteratur. Retrieved from <http://www.mt-archive.info/NLPCS-2011-Carl-1.pdf>.
- Gibbs, R. W. (2010). The dynamic complexities of metaphor interpretation. *D.E.L.T.A.*, 26, special issue, 657–677.
- Gibbs, R. W., & Tendahl, M. (2008). Complementary perspectives on metaphor: Cognitive linguistics and relevance theory. *Journal of Pragmatics*, 40, 1823–1864.
- Gibbs, R. W., & Tendahl, M. (2011). Coupling of metaphoric cognition and communication: A reply to Deirdre Wilson. *Intercultural Pragmatics*, 8(4), 601–609.
- Gibbs, R., Tendahl, M., & Okonski, L. (2011). Inferring Pragmatic Messages from Metaphor. *Lodz Papers in Pragmatics*, 7(1), 3–28.
- Green, S., Heer, J., & Manning, C. D. (2013). The Efficacy of Human Post-Editing for Language Translation. ACM Human Factors in Computing Systems (CHI). Retrieved from <http://vis.stanford.edu/papers/post-editing>
- Gutt, E. (1992). *Relevance Theory: a guide to successful communication in translation*. New York/Dallas: Summer Institute of Linguistics and United Bible Societies.
- Krings, H. (2001). *Repairing Texts: Empirical Investigations of Machine Translation Post-Editing Processes*. Kent, OH: The Kent State University Press.
- O'Brien, S. (2006a). *Machine-Translatability and Post-Editing Effort: an empirical study using Translog and Choice Network Analysis*. (Ph.D. thesis, Dublin City

- University, Ireland).
- O'Brien, S. (2006b). Pauses as indicators of cognitive effort in post-editing machine translation output. *Across Languages and Cultures, Budapest, 1*(7), 1–21.
- O'Brien, S. (2007). An empirical investigation of temporal and technical post-editing effort. *Translation and Interpreting Studies (TIS)*, 2(1), 83–136.
- O'Brien, S. (2009). Eye tracking in translation process research: methodological challenges and solutions. In I. M. Mees, F. Alves, & S. Gopferich (Eds.). *Methodology, technology and innovation in translation process research: a tribute to Arnt Lykke Jakobsen* (pp. 251–266). Copenhagen studies in language, 38. Copenhagen: Samfundslitteratur.
- Sperber, D., & Wilson, D. (1986). *Relevance: Communication and cognition*. Oxford: Blackwell.
- Sperber, D., & D. Wilson. (2008). A deflationary account of metaphors. In R. W. Gibbs (Ed.), *The Cambridge Handbook of Metaphor and Thought* (pp. 84–108). New York: Cambridge University Press.
- Steen, G., Dorst, A. G., Herrmann, B., Kaal, A., Krennmayr, T., & Pasma, T. (2010). *A method for linguistic metaphor identification: from MIP to MIPVU*. Amsterdam/Philadelphia: John Benjamins.
- Unger, C. (2001). *On the cognitive role of genre: a relevance-theoretic perspective*. (Ph.D. dissertation, University of London, UK).

Appendix A: Source Text

The Tea Party Pork Binge

They brought the nation to the brink of default over spending, but a Newsweek investigation shows Tea Party lawmakers grabbing billions from the government trough. Plus, view the letters submitted by the 'Dirty Dozen'. House Majority Leader Eric Cantor, the Republican leadership's tether to the Tea Party, flutters the hearts of the government-bashing, budget-slicing faithful with his relentless attacks on runaway federal spending. To Cantor, an \$8 billion high-speed rail connecting Las Vegas to Disneyland is wasteful "pork-barrel spending." The Virginia Republican set up the "You Cut" Web site to demonstrate how easy it is to slash government programs. And he made the Department of Housing and Urban Development the poster child for waste when he disclosed that the agency was paying for housing for Ph.D.s. But away from the cameras, Cantor sometimes pulls right up to the spending trough, including the very stimulus law he panned in public. [...]

As the government showdown over debt continues — the so-called congressional supercommittee negotiating cuts has been floundering for weeks — Newsweek found about five dozen of the most fiscally conservative Republicans, from Tea Party freshmen like Allen West to anti-spending presidential candidates like Rick Perry and Ron Paul, trying to gobble up the very largesse they publicly disown, in the time-honored, budget-busting tradition of bringing home the bacon for local constituents.

Appendix B: Google Translate output

O Tea Binge Pork Partido

Eles trouxeram a nação à beira da inadimplência sobre os gastos, mas uma investigação Newsweek mostra legisladores Tea Party agarrando bilhões da calha do governo. Além disso, visualizar as cartas apresentadas pela 'Dirty Dozen'.

Casa Líder da Maioria Eric Cantor, amarrar a liderança republicana para o Tea Party, palpita o coração do governo-bashing, orçamento-corte fiel com seus ataques implacáveis sobre os gastos federais em fuga. Para Cantor, 8 bilhões de dólares ferroviária de alta velocidade ligando Las Vegas a Disneyland é um desperdício "Os gastos de porco barril." O Republicano da Virgínia criou o "Você Cut" site para demonstrar como é fácil de cortar programas de governo. E ele fez o Departamento de Habitação e Desenvolvimento Urbano a criação do poster para os resíduos, quando ele revelou que a agência estava pagando por habitação para doutores. Mas longe das câmeras, Cantor, por vezes, puxa até o vale de gastos, incluindo a lei de estímulo muito, ele criticou em público. [...]

Como o confronto do governo sobre a dívida continua a supercommittee-so-called do Congresso negociando cortes foi tropeçando por semana-Newsweek encontrados cerca de cinco dezenas dos republicanos mais conservadores fiscais, a partir de calouros Tea Party como Allen West para anti-gastos candidatos presidenciais como Rick Perry e Ron Paul, tentando engolir a generosidade muito que repudiar publicamente, no time-honored, tradição orçamento-rebentando de trazer para casa o bacon para constituintes local.

Appendix C: Systran output

O frenesi da carne de porco do tea party

Trouxeram a nação ao limiar do defeito sobre a despesa, mas os legisladores de um tea party das mostras da investigação de Newsweek que agarram bilhões da calha do governo. O sinal de adição, vê as letras submetidas “pela dúzia suja.”

Abrigue o cantor de Eric do líder da maioria, o barão da liderança republicana ao tea party, vibrações os corações do governo-bashing, orçamento-corte fiel com seus ataques implacáveis na despesa federal do fugitivo. Ao cantor, um trilho \$8 bilhões de alta velocidade que conecta Las Vegas a Disneylândia é do “despesa desperdiçada carne de porco-tambor.” A Virgínia que o republicano estabelece “você cortou” o Web site para demonstrar como fácil é reduzir programas governamentais. E fez ao departamento de habitação e desenvolvimento urbano a criança do cartaz para o desperdício quando divulgou que a agência estava pagando abrigo para Ph.D.s. Mas longe das câmeras, o cantor puxa às vezes até à calha da despesa, incluindo a lei que mesma do estímulo filtrou em público. [...]

Enquanto o governo que a prova final sobre o débito continua- cortes de negócio do supercommittee do congresso assim chamado tem chafurdado para semana-Newsweek encontrou aproximadamente cinco dúzias dos republicanos o mais fiscal conservadores, dos caloiros do tea party como Allen ocidental aos candidatos presidenciais da anti-despesa como Rick Perry e Ron Paul, tentando devorar acima da largueza mesma repudiam publicamente, na tradição tradicional, orçamento-rebentando de trazer em casa o bacon para componentes locais.