The International Journal for Translation & Interpreting Research
trans-int.org

An investigation of the relevance-theoretic approach to cognitive effort in translation and the post-editing process

Kyoko Sekino
Federal University of Minas Gerais
kyokosekino@ufmg.br

DOI: ti.106201.2015.a09

Abstract: The present study presents findings from an experiment that investigates the cognitive effort of eight translators and eight students during translation and post-editing (PE) tasks from Japanese to Brazilian Portuguese. Relevance theory is used as a theoretical framework to contextualise the object of study—namely conceptual and procedural encodings—and interprets the use of these encodings as the exertion of cognitive effort by participants to infer the intended message. This study replicates Alves & Gonçalves's (2013) methodology and applies it not only to translation but also to post-editing. The results show that participants mainly edit procedural encodings. However, post-editing appears to require participants to exert greater effort since both encodings are edited. No time difference is observed between the two tasks.

Keywords: relevance theory, conceptual encoding, procedural encoding, post-editing, cognitive effort

1. Introduction

In translation, the main goal is to produce a target text that is as meaningful as the source text. Relevance theory, first introduced by Sperber & Wilson (1986), describes human linguistic communication as being ostensive in nature. The authors recognise that communication simultaneously requires information to be encoded and decoded and for the parties to infer meaning from the communicative act. The notion of inference is the major point of discussion in relevance theory that reveals our cognitive commitment during communication. Moreover, relevance theory asserts that people express more concepts than what can be explicitly encoded in linguistic expression, and as a result, inferential meaning is used to enrich the communicative act to access the sender's intention. The same can be said of bilingual communication and scholars such as Gutt (1991/ 2000) have argued that relevance theory encompasses translation, insofar as readers of a translated text can access the intention of the author through inference. The same ostensive-inferential communication can also be extended to post-editing (PE) as the resulting text is a translation. In this study, we compare the cognitive effort exerted to render a target language version of a text during the translation and post-editing tasks. We employ a relevance-theoretic approach and observe participants' revisions of previously translated parts. A total of sixteen (16) participants, eight translators and eight students of Japanese, performed either translation or post-editing tasks from Japanese (L2) to Brazilian Portuguese (L1). For the post-editing task, Google Translate provided the machine translation output.

2. Theoretical underpinnings

In the last few decades, research in translation studies has examined the translation process as well as the translation product and has incorporated research methodologies that are theoretically grounded in other fields such as psychology. These methods are used in an effort to better understand translators' cognitive processes. Several data collection methods have been employed to map cognitive processes in translation studies such as thinking-aloud protocols, keystroke logging, and eye tracking.

Jakobsen (1999), for example, developed Translog as an instrument that records each keystroke and mouse movement during the translation process. An important feature of Translog is the ability to record the translation process in real-time such that it can be replayed for analysis. This approach to collecting data has opened up new directions and inspired further methodological innovation to access real-time process data. In the subsequent version, Translog 2006, several new options became available for data collection and analysis. One of the variables in the translation process that Translog has made observable is the real-time translation unit (TU). According to Malmkjaer (1998), a process-oriented TU is a "stretch of the source text that the translator keeps in mind at any one time, in order to produce translation equivalents in the text he or she is creating" (p. 286). Dragsted (2004, 2005), in her study of cognitive segmentation, addresses the notion of the translation unit within the context of cognitive translation processes. Her conception of a TU aligns with Malmkjaer's process-oriented TU, and operationalises the concept based on cognitive resource allocation and memory. Dragsted considers the translation unit as the translator's focus of attention at any given moment during the translation process. This attentional focus can vary and depends largely on the translator's viewpoint.

Alves & Vale (2009, 2011) build on Dragsted's assumption and incorporate Jakobsen's (2005) findings on how to observe peak translator performance. Alves & Vale (2009, 2011) propose a classification of TUs into micro and macro TUs, and segment their Translog log file using Jakobsen's (2005) proposed 2.4 second pause. The authors define a micro TU as "the flow of continuous TT production – which may incorporate the continuous reading of ST and TT segments – separated by pauses during the translation process as registered by key-logging and/or eye-tracking software" (Alves & Vale, 2009, p. 257). In the same vein, the authors define a macro TU as "a collection of micro TUs that comprise all the interim text productions that follow the translator's focus on the same ST segment from the first tentative rendering to the final output that appears in the TT" (ibid). Therefore, a macro TU incorporates all the text production segments during the process, starting from the initial focus of attention that triggered a given micro TU. In order to analyse process data recorded in the Translog log file, the text products must first be segmented into micro TUs using pauses. Then, macro TUs need to be identified that manifest recursive movements. These movements are changes to already rendered micro TUs in the text production. This process—segmentation by pauses followed by the construction of macro TUs—highlights instances in which translators exert additional effort to produce the final text.

Alves & Gonçalves (2007, 2013), have used micro and macro TUs to study relevance theory hypotheses empirically; these studies have been possible due to recent methodological innovations in observing translators' behaviour. One particular behaviour that can now be observed is editing—the effort to produce a final text in order to facilitate readers' inferencing by providing adequate contextual clues.

Sperber & Wilson (1986/1995) hypothesise in relevance theory that an external stimulus, such as utterance in verbal communication, can be manifested in linguistic representations. The authors argue that during this linguistic manifestation, the mental representation of the speaker is retained as a higher-order representation which is, in fact, the real message of that stimulus. The term *representation* is used in this context to refer to anything that has meaning, like maps and pictures, as opposed to just being itself.

Within the relevance theory framework people perform distinct processes during linguistic communication, namely encoding/decoding and inference. Relevance theory postulates that, in linguistic communication, a message that has been encoded using lexical clues is first decoded and then enriched by inferences. These inferences consider the contextual information provided by the speaker to access his or her intention or real message. For example, in an utterance such as, *My steak is raw* (Higashimori & Yoshimura, 2003, p. 51), the meaning of the utterance decoded lexically is the steak of the person referred to as "I (My)" is raw and has not been cooked. Suppose this utterance occurred in a steak house when that steak was served on the table (contextual information), the meaning of the utterance would be interpreted as *my steak is not well-done* or *not cooked enough, as I expected*. This interpretation expected by the speaker, in conjunction with the context, is made by *inference*. Sperber & Wilson (1986) claim that the sender of the message transmits it to the receptor in order for the message to be understood (cognitive benefit) with the minimum cognitive effort (processing cost) exerted.[1] In other words, people seek to optimise the cognitive system, by using minimally sufficient lexical clues, in order to avoid exerting unnecessary processing effort—once the receptor comprehends a message as expected by the speaker, the former stops further processing. Therefore, as Moeschler (2002) explains, RT presents an equilibrium between cognitive effort and its corresponding generated effects (contextual effects).The author describes:

> [i]n other words, economy is a property of the cognitive systems devoted to utterance interpretation, as well as required, in order to insure a successful communication, by the computational devices, which combine with linguistic decoding to yield the intended meaning of the utterance. (Moeschler, 2002, p. 1)

With regard to encodings, the distinction between conceptual and procedural encodings (CC and PC) was first introduced by Blakemore (1987), and discussed further by Wilson & Sperber (1993), Moeschler (1998), and Blakemore (2002). According to Moeschler (2002), people communicate events, states and about relations between these events, by adding, modifying, and replacing information. In order to update and (re)organise mental representation of these events, information is needed on how to manipulate these representations, namely conceptual information and procedural information. Conceptual information is related to the concepts that linguistic expressions represent, while procedural information is how to manipulate mental representations accessed by concepts. Moeschler (2002) considers natural languages should have specialised some of their expressions for these functions, as long as they are efficient, and postulates that "lexical categories map onto conceptual information and functional categories map onto procedural information" (p. 5). Blakemore (2002) explains that conceptual encodings consist of content words such as nouns, adjectives, verbs, while procedural ones are instructions such as indicators of tense, discourse connectives, word orders and so on.

[1] This is termed the 'cognitive principle of relevance' by Sperber & Wilson (1986/1995).

Wilson & Sperber (1993), Blakemore (2002), and Wilson (2011), also claim that procedural encodings sometimes entail more effort on the part of receptors than conceptual ones, in order to infer the true meaning of what the sender intended to transmit.

Gutt (1991/2000), in turn, applies relevance theory to translation process research. Understood using an inferential approach to communication study, the author recognises the proposition that whole efforts in communication originate in the cognitive core or cognitive system. At the same time, however, effort involves inference from ostensive linguistic clues, which differs from mere encoding and decoding of utterances. In terms of relevance theory:

> [t]ranslation is viewed as a higher-order act of communication (HOAC) - an act of communication that is about another (lower-order) act of communication. Any act of (ostensive) communication necessarily involves two focal elements: the stimulus, which is the perceptible element, e.g. an utterance or text in verbal communication, and the interpretation, the body of thoughts which the communicator intends to share with others. (Gutt, 2005, p. 25)

Gutt affirms that in the case of translation, relevance theory is applicable, because the translation task is by nature a creation of interpretative similarity — i.e., an interpretive translation that shares a resemblance with the original. This author adds that the empirical observation of these human efforts with respect to inference requires adequate methods that enable us to observe instances of cognitive commitment involved during translation.

Drawing on Gutt's (1991/2000, 2005) work, recent studies by Szpak (2012) and Alves & Gonçalves (2013) have investigated translators' cognitive effort. Translators' behaviour that indicates a return to previously translated material and its subsequent modification can be one way to observe how much effort was exerted. Using instruments such as Translog and a Tobii T-60 eye tracker, and retrospective verbal protocols, Alves & Gonçalves count the number of edits made in conceptual encodings (CC) and in procedural encodings (PC) during the translation process.

To annotate these revisions, Alves & Gonçalves establish a taxonomy to classify edits made in macro TUs. In this study two classes of annotations are used, using the following tags: for translation phases annotation are made with tags [P0], [P1], [P2] and [P3], which indicate when or in which translation phase edits may occur. These tags are in line with the phases defined by Jakobsen (2003), who identifies three translation phases: orientation, draft, and revision. [P1] corresponds to edits during the draft phase; [P2] for edits during the revision phase; [P3] for edits in both draft and revision phases. [P0], in turn, is a special tag that indicates text production of translation during the draft phase that is not subsequently modified in the process, differentiating [P0] from [P1]. Linguistic edits are annotated with tags such as [l], [m], [p], [t] or [c], in order to record the conceptual and procedural encodings: [l] for edits in content words such as verbs, adjectives, nouns, etc.; [m] for morphosyntactic indicators such as discourse connectives, tense, indexical referents, etc.; [p] for structural modifications, e.g., from cleft clause to noun phrase; [t] for typographical mistakes; and, [c] for the completion of a word.[2]

The present study partially replicates Alves & Gonçalves' methodology of data collection and analysis in a different linguistic pair, and includes post-editing

[2] If one word was corrected several times for different reasons — for example, if a typo was corrected and soon gender was also corrected — annotation tags [t] and [m] would both be attributed.

of machine translation output along with the human translation task. Some early observations about the practice of post-editing have been reported by Vasconcellos (1986a, 1986b), Somers (1997), Hutchins (1998), Senez (1998), Allen (2001; 2003) and O'Brien (2002), among others; however, Krings (2001) presented the first systematic study of post-editing, in which three types of PE effort were identified: temporal, cognitive and technical. Krings claims that of these, temporal effort can be interpreted as the sum of cognitive and technical efforts. Based on his claim, temporal and technical effort in PE have been measured in other studies with use of key-loggers, as mentioned previously.[3] O'Brien (2006), for example, uses Translog and Choice Network Analysis (Campbell, 2000) to analyse pauses in PE. O'Brien used a pause ratio to measure cognitive effort, and calculated this measure by dividing total time in pause by the total time in segment to explore the relationship between the pause ratio and segments that may contain negative translatability indicators[4] (NTIs) of the raw machine output. Almeida & O'Brien (2010), using readability metrics, investigate the PE process of participants and includes participants who have almost no experience of PE task. Tatsumi (2011) conducts a PE experiment with the Japanese-English language pair in order to identify types of post-editing.

3. Experimental design and method of analysis

The present study's methodology is based on earlier studies conducted at the Laboratory of Experimentation for Translation (LETRA/UFMG), such as Alves & Vale (2009; 2011), Szpak (2012), and Alves & Gonçalves (2013). These studies are characterised by the triangulation of data collection methods, incorporating quantitative methods such as keystroke logging and eye tracking with qualitative methods such as retrospective verbal protocols. The principal objective of this study is to identify which type of encodings—conceptual or procedural—demands more cognitive effort of participants when working between Japanese and Brazilian Portuguese. In line with Alves & Gonçalves (2013), we consider that editing conceptual and procedural encodings is closely related to the translators' cognitive effort to produce contextual effects in final translation text as well as the post-edited text. Therefore, we understand that a high number of edits on procedural encodings is an indication of translator/post-editor's cognitive commitment associated with inference.

We also aim to verify whether the methodology originally created to observe the translation process is appropriate to study post-editing, and if the taxonomy developed for translation fits PE analysis.

Brazilian Portuguese professional translators and Japanese language students either translated or post-edited one Japanese text into Brazilian Portuguese (L1 of all the participants). The source text (ST) in this study was extracted from an IT-related website[5] and consists of 281 Japanese characters. For the post-editing task,

[3] Several keystroke logging computer programs exist and have been used in writing research. One such program, InputLog (cf. Leijten & Waes, 2013), has also been used in translation process studies, for example by Ehrensberger-Dow & Perrin (2013). For more information, see www.inputlog.net/ .

[4] The indicators of difficult translation segments such as passive voice was initially introduced by Underwood & Jongejan (2001) as 'translatability indicator', O'Brien (2005) prefers using the term 'negative translatability indicator' or NTIs, to emphasize a 'problematic' linguistic feature in the source text that poses challenges for machine translation systems.

[5] www.icr.co.jp, accessed in April 2012.

Google Translate was used as the machine translation system (MT) to provide the text used in the post-editing task. The decision to use Google Translate was made in the light of the limited options available to provide an L1 translation in this language pair. Each participant either directly translated from the Japanese ST, or post-edited the machine output in consultation with the Japanese source text. No time limit was imposed and participants were allowed to use external online resources.

As no other study has investigated this language combination, we adopted the hypotheses of Alves & Gonçalves (2013) as our working hypotheses for both tasks. First, we hypothesise that procedural encodings will be modified more than conceptual encodings in macro TUs. Second, since Google Translate is supposed to assist participants in constructing a final translation text, the number of edits in conceptual encodings and procedural encodings in PE will be less than that in translation. Third, participants in the post-editing task will complete their task more quickly than those completing the translation task.

For the data collection, all the tasks were recorded using keystroke –logging using Translog 2006. After each task, participants were asked to report their translation or PE process while watching the Translog replay. This retrospective verbal protocol was recorded; however, the results from these recordings are not presented here.

With respect to data analysis, the data were first recorded in the Translog log file and then specific thresholds for pause value were established. We replicated Alves & Gonçalves's (2013) study in that we adopt a 2.4 second pause value for both translation and PE analysis. This decision is in light of the exploratory nature of this study. The target text was then segmented by pauses, in order to observe each participant's processing behaviour, forming first micro TUs, and later, macro TUs. Once the micro TUs had been segmented, all the TUs were annotated with the tags established by Alves & Gonçalves's taxonomy.

Finally, the variables that we investigate are the number of macro TUs, the number of modifications made in conceptual encodings as well as in procedural encodings, and the total pause time to answer our hypothesis.

4. Results

Data were collected from March to May in 2013. The valid protocols comprise data from 4 translators for the translation task, 4 translators for the PE task, 5 students for the translation task and 3 students for the PE task, for a total of 16 participants. In this section, we first report the number of macro TUs created by the participants according to their tasks in order to reveal differences between these tasks. Next, we outline the number of edits made in CC and PC and first compare translators in the translation task and the PE task; the same is done between students in the both tasks. Following, the results are presented on CC and PC edits within the translation task, comparing translators and students. This comparison is also made within the post-editing task. Finally, we present findings with regard to time and pause recorded in the tasks.

We conducted the translation/PE phase analysis first, that is, the number of [P0], [P1], [P2] and [P3] translation units. We compared the translation task group and PE task group. This comparison using the Wilcoxon rank sum[6] test showed no

[6] All non-parametric tests are Wilcoxon rank sum tests, unless otherwise noted.

significant difference.[7] There seems to be a pattern of a greater number of edits in the draft phase and a very small number of edits in the revision phase: an average of 92.6 percent of the edits were realised in the P0/P1 draft stage (for translation: M = 28.22, SD = 11.80; for PE: M = 37.71, SD = 10.13). In other words, only a few more than 7 percent of modifications were made in revision phase in the tasks (Figure 1). We understand, from this result, that most of the participants are satisfied with their rendering or PE worked out during the draft phase. We recognise the possibility of a possible white-coat effect; however, the collected data cannot clarify this point.

As mentioned earlier, [P0] is a special tag attributed to micro TUs created in the translation draft phase like [P1] but that does not undergo any modification during the whole translation process. We observe that TUs annotated with the [P0] tag represent 45.2 percent of all TUs (Figure 2). From this observation, we assume that the participants engaged in the translation task were satisfied with their original choices and that these required no further modifications. This proportion increases if we take into account the instances of typographical mistakes.

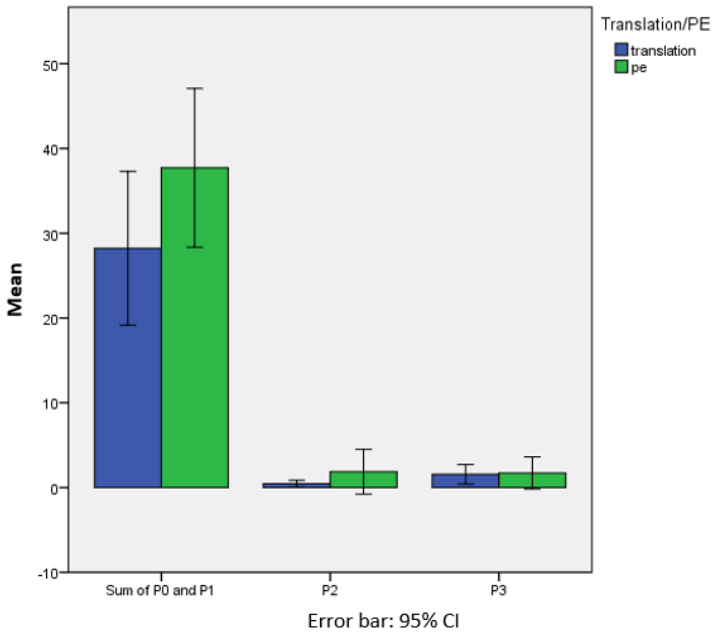


Figure 1: The mean of number of macro TUs in the draft phase and revision phase in translation and PE

For the number of CC and PC edited in the both tasks, there was a statistically significant difference for both translators and students. For the translators (4 engaged in translation task and 4 in PE task) the results show that CC edits during the post-editing task exceed those made during the translation task, z = -2.32, p = 0.01. Likewise, a greater number of PC edits were made during the post-editing, z = -2.31, p = 0.01. Thus translators when post-editing made more changes for both CC and PC more than they did during in the translation task. The same result was obtained among student group, such that students in PE task also edited a greater number of CC (z = -2.24, p = 0.013) and PC (z = -2.24, p = 0.013) in the post-editing task than in translation. The overall comparison regardless of the participants categories (9 in translation and 7 in PE task) the

[7] For the nature of PE which does not have P0, P1 in PE is considered equivalent with P0 and P1 in translation.

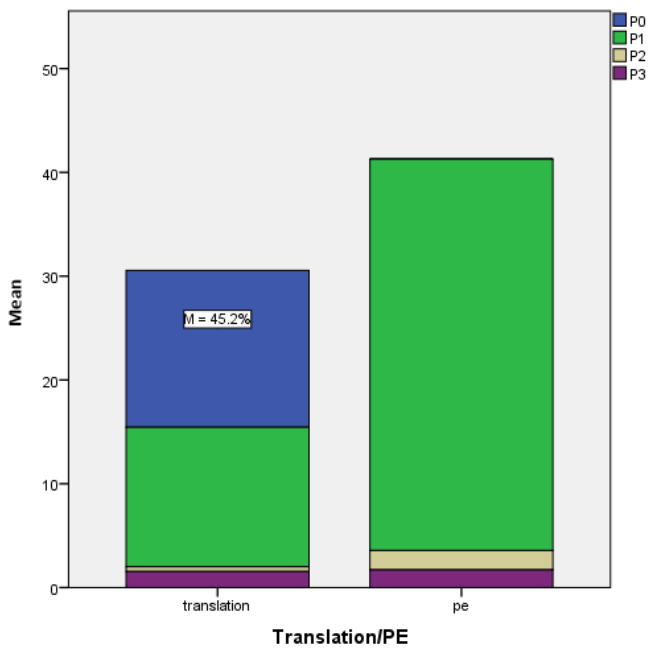
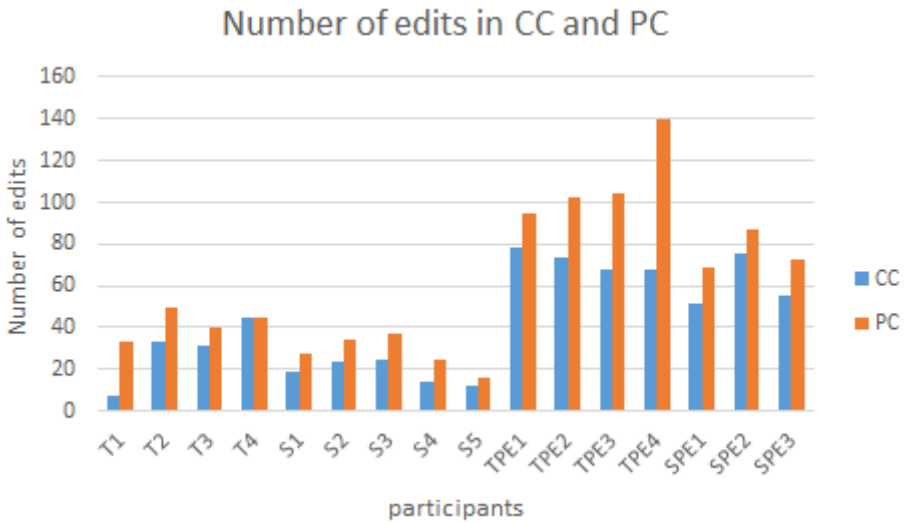


Figure 2: The mean of the number of macro TUs, P0, P1, P2 and P3 in translation and PE.



(T: Professional translator in the translation task, S: Student in the translation task, TPE: Translator in the PE task, SPE: Students in the PE task)

Figure 3: CC and PC individual edits.

same was observed: the comparison on CC in both tasks, for CC edits, $z = -3.34$, $p < 0.01$; for PC edits, $z = -3.33$, $p < 0.01$.These results indicate that post-editing demands greater effort than translation as the number of edits in CC and PC are greater in PE task.

Within the translation task (4 translators and 5 students), there was no significant difference found between the translators and the students in either CC or PC edits. The results were the same for the PE task (4 translators and 3

students) as for translation. These results imply that there is no professional characteristics identified in this comparison.

With regard to the total time spent for each task, the result suggest that there is no significant difference between two tasks. This result is somewhat unexpected in light of the PE task, since the provision of the MT output is to assist the participants in the production of a final translation text. We also compared the production time within the groups and found no difference.

We further analysed time-related data by calculating the proportion of sum of pause time in total production time. We found a significant difference, $t(14) = -2.90$, $p = .012$. This result indicates that the participants paused more during the PE task than during translation. These pauses might be directly associated with reading. This hypothesis is somewhat intuitive, since in the case of bilingual post-editing, post-editors necessarily have more material to read than translation.

5. Discussion

The present study's principal aim was to investigate whether the number of edits in procedural encodings (PC) are more numerous than those made in conceptual encodings (CC). The study's results did, in fact, reveal a greater number of edits in PC. From a relevance-theoretic view, people may make more effort to process PC for successful communication because they are responsible for procedural information closely related with inferences needed to manipulate the relation between events. It was not, however, possible to support the study's other hypotheses, such as post-editing task being less effortful than translation. First, no temporal difference is observed between the tasks; second, there are significant number of edits in PC observed in comparison with that of translation. One might question the quality of the raw machine output for the PE task, as well as of the participants' performance of the PE task. Nevertheless, overall tendencies may be considered, acknowledging this study to be an exploratory one and may serve as a starting point for further research.

First of all, with regard to post-editing, its purpose is to reduce translation text production time and cost through the use of automatic machine translation systems as an integrated part of translation (Allen, 2001). In such instances of machine-human interaction, machine output must be of adequate quality in order for it to relieve human work. In the linguistic pair used in the present study, the demand for translations is quite low with respect to volume; consequently, no up-to-date online dictionary exists. Google Translate is a rather convenient tool, free of charge, and frequently used as an online dictionary as well as an automatic MT system. All the participants in the translation task could consult Google Translate as an external resource. Surprisingly, however, the results show that the PE task required as much time as the translation. There may be several reasons for these results: first, the participants' unfamiliarity with the task; second, problems with the raw machine output quality. The ST used was randomly selected from within the journalistic genre and did not use controlled authoring; nor was the normal procedure for post-editing followed, such as pre-editing or pre-processing of the raw MT text (Krings, 2001; Allen, 2003) followed.

One reason for the increased time spent in target text production during the PE task may have been the sentence structure problem in the Google raw output; Japanese and Brazilian Portuguese differ in sentence structure substantially. In Japanese, words, phrase and clauses that modify nouns or phrasal nouns, as well as relative clauses, are prepositional, instead of post-positional (Shibatani, 2009). According to the TAUS Evaluating Post-Editor Performance

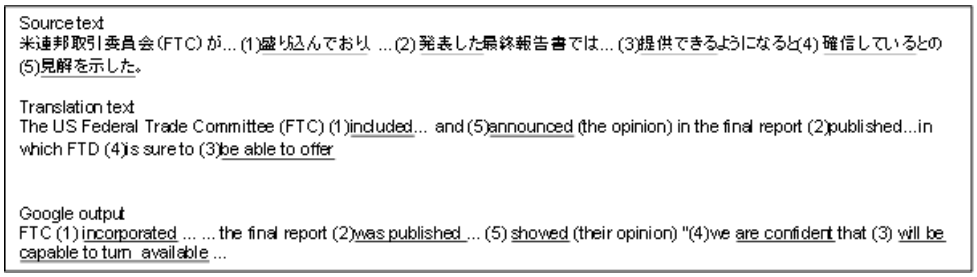


Figure 4: Comparison of the second sentence's main verbs.

Guidelines (2014), word order persists as one of the problem areas. In the PE task, for example, the second sentence certainly does illustrate this problem.

As seen in Figure 4, the difficulty post-editors faced in the PE task may have been in reorganising the original word order observed in the five principal verbs in the sentence: original verb word order (1)-(2)-(3)-(4)-(5) to Portuguese word order (1)-(5)-(2)-(4)-(3), from the intervening sequence Google generated, (1)-(2)-(5)-(4)-(3). Furthermore, there are several other word order problems in the second sentence, such as those involving phrasal nouns and relative clauses, as well as tense, plurality, and above all, the switch from S-O-V structure to S-V-O, etc. Nagatsuji (2012), for example, explains that the type of verb connection, as observed in verb (1) in the ST, has no tense marker and expresses a "cause" in the cause-effect relationship. However, the ST is quite ambiguous when you look to find the "effect" among the other verbs (2), (3), (4) and (5). Needless to say, in the machine output the verbs are quite confusing in terms of grammatical subject, tense and connection with other part of the discourse. These create problems in procedural encodings to rebuild a meaningful text. These factors can be considered to have contributed to the delay in PE, which otherwise, would have be expected to be much faster than human translation.

With regard to the professional performance, our results show no difference between translators and students in each task. More specifically, we expected to observe professional behaviour such as distinct problem-solving behaviour from the translators in PE, even if none of them had previous experience. However, it appears that the PE task entailed difficulties similar to those found in translation. Consequently, we were unable to identify a distinct effect of the task itself. Although Almeida & O'Brien (2010) predict that professional translators, including those who have no experience in PE, would handle PE better than non-experienced participants, our findings do not support this hypothesis.

When we examine revisions in the post-editing task, edits in procedural encodings occur more than in conceptual ones, and in both groups. Once again, this result may be related to the unrefined machine output that results from two distant languages as well as from task unfamiliarity. Participants often deleted and rewrote word groupings proposed by Google Translate. In fact, rewriting often resulted from the relocation of word groups, a corrective action presumably preferred by participants when copying and pasting. However, these keystrokes, as well as copy-paste events, were not to reconstruct words or chunks exactly the same as before but to rewrite them in a slightly changed in order, tense, plurality etc. Specifically, Japanese does not have articles, gender or plurality except for certain words (Shibatani, 2009), so procedural encodings needed to be edited, for the sake of Portuguese agreement. The above-mentioned facts may have been one of the reasons that the data show a significantly higher number of edits on PC than in CC. Even editing CC by replacing words in Portuguese requires post-editors to

commit to an edit and supply the surrounding PC elements. Correction of word order requires subsequent changes in the discourse connectives especially to reorganise the sentence structure. Therefore, as Krings (2001) indicates, there are two layers of cognitive effort required in PE. One is at a lower level, at which post-editors mainly seek linguistic correctness. As already pointed out with regard to our data, a share of the large number of PC edits may be closely related to these lower level efforts. The other level, that is the upper level, involves the post-editor's engagement with the discourse level. In this study, because the raw machine output had not been adequately refined to permit observation of discourse level modifications in CC and PC, it would be problematic to conclude that more PC modification was observed due to improvement at the discourse level.

As postulated by Wilson & Sperber (1993), Blackmore (2002), and Wilson (2011) and investigated by Alves & Gonçalves (2013) within the field of translation process research, PC may demand as much, or more, cognitive effort than CC. With regard to the number of edits, our data present PC is more predominant than CC, consequently, we may say that the participants have exerted more effort by editing PC than CC. In Moeschler's (2002) terms, the participants were engaged with more procedural categories to work out the manipulation of procedural information.

Finally, our first working hypothesis that edits in PC are more predominant than those of CC in macro TUs is supported by this study. The second hypothesis, however, that post-editing assists translators and post-editors in constructing a final translation text was not confirmed, as edits in PC and CC during the PE task are greater than those made during translation. For the third hypothesis, as the temporal effort in PE seemed to be similar to that of translation, we are unable to confirm in this study. From the second and third hypotheses, we may conclude that PE task required more effort from the participants in this study; however, further investigation is needed in light of some of the discussed limitations.

6. Conclusion

The study confirms that the Alves & Gonçalves (2013) method is a promising contribution to translation and PE process research for the Japanese-Brazilian Portuguese language pair. We would highlight the fact that the relevance-theoretic approach, combined with observation of effort indicators such as time, pause and number of editions made in CC and CP, may reveal the allocation of cognitive effort during the tasks. However, we admit that certain methodological issues should be taken into account, such as the quality of the raw machine output, the instructions given before the PE task is performed, the of external resources to participants during the tasks, and the sample size, and the influence that training sessions for participants may play, among others. These are all areas that could serve as the foundation for future research projects. Most importantly, we find the annotation of PE data using the Alves & Gonçalves (2013) taxonomy to be quite complex—which implies a need to improve the mechanical tools. Such improvements in future investigations may enable researchers to be somewhat more successful in data analysis and the statistical results obtained.

About the author

Kyoko Sekino is a PhD student at Federal University of Minas Gerais (UFMG) in Brazil. She belongs to the Laboratory for Experimentation in Translation (LETRA) and is mainly engaged in translation process research. She also teaches Japanese at University of Brasilia, Department of Foreign Languages and Translation in Brazil.

References

Allen, J. (2001). Postediting: an integrated part of translation soft program. *Language International*, April, 26–29.

Allen, J. (2003). Post-editing. In H. S. Somers (Ed.), *Computers and translation: a translator's guide* (pp. 297 – 317). Amsterdam: John Benjamins.

Almeida, G., & O'Brien, S. (2010). Analysing post-editing performance: correlations with years of translation experience. *14th Annual Conference of the European Association for Machine Translation.* Saint Rafaël.

Alves, F., & Gonçalves, J. L. (2007). Modelling translators' competence: Relevance and expertise under scrutiny. In Y. Gambier, M. Shlestinger, & R. Stolze (Eds.), *Doubts and directions in translation studies: Selected contributions from the EST Congress, Lisbon 2004* (pp. 41–55). Amsterdam: John Benjamins.

Alves, F., & Gonçalves, J. L. (2013). Investigating the conceptual-procedural distinction in the translation process: a relevance-theoretic analysis of micro and macro translation units. *Target*, *25*(1), 107–124.

Alves, F., & Vale, D. C. (2009). Probing the unit of translation in time aspects of the design and development of a web application for storing, annotating and quering translation process data. *Across Languages and Cultures*, *10*(2), 251–273.

Alves, F., & Vale, D. C. (2011). On drafting and revision in translation: a corpus linguistics oriented analysis of translation process data. *Translation: Computation, Corpora, Cognition*, *1*(1), 105–122.

Blakemore, D. (1987). *Semantic constraints on relevance*. Oxford: Basil Blackwell.

Blakemore, D. (2002). *Relevance and linguistic meanings: The semantics and pragmatics discourse makers*. Cambridge: Cambridge University Press.

Campbell, S. (2000). Choice Network Analysis in translation research. In M. Olohan (Ed.) *Intercultural faultilines: Research models in translation studies: Textual and cognitive aspects* (pp. 29–42). Manchester: St. Jerome.

Dragsted, B. (2004). *Segmentation in translation and translation memory systems: An empirical investigation of cognitive segmentation and effects of integrating a TM system into the translation process*. København: Samfundslitteratur. (Ph.D. series; No. 5, Vol. 2004). Retrieved from http://www.cbs.dk/en/research/cbs-publications.

Dragsted, B. (2005). Segmentation in translation: Differences across levels of expertise. *Target*, *17*(1), 49–70.

Ehrensberger-Dow, M., & Perrin, D. (2013). Applying a newswriting research approach. *Target*, *25*(1), 77–92.

Gutt, E. A. (1991/2000). *Translation and relevance: Cognition and context* . Manchester: St. Jerome.

Gutt, E. A. (2005). On the significance of the cognitive core of translation. *The Translator*, *11*(1), 25–49.

Higashimori, I., & Yoshimura, A. (2003). *New development of relevance theory: cognition and communication*. Tokyo: Kenkyusha.

Hutchins, J. (1998). Twenty years of translating and the computer. *Translating and the Computer 20: Proceedings of the Twentieth International Conference on Translating and the Computer.* London: Aslib.

Jakobsen, A. L. (1999). Logging target text production by Translog. *Copenhagen Studies in Language*, *24*, 9–20.

Jakobsen, A. L. (2003). Effects of think aloud on translation speed, revision and segmentation. In F. Alves (Ed.), *Triangulating translation. Perspectives in process oriented research*. (pp. 69–95.) Amsterdam: John Benjamins.

Jakobsen, A. L. (2005). Instances of peak performance in translation. *Lebende Sprachen*, *3*, 111–116.

Krings, H. P. (2001). *Repairing texts: Empirical investigations of machine translation post-editing processes.* (G. S. Koby, Ed.) Kent, OH: The Kent University Press.

Leijten, M., & Waes, L. v. (2013). Keystroke Logging in Writing Research: Using Inputlog to Analyze and Visualize Writing Processes. *Written Communication, 30*(3), 358–392.

Malmkjaer, K. (1998). Unit of Translation. In M. Baker (Ed.) *Routledge Encyclopedia of Translation Studies* (pp. 286–88). London: Routledge.

Moeschler, J. (1998). Directional inferences and the conceptual/procedural encoding distinction. *Proceedings ot the Relevance Theory Workshop, University of Luton (8–10 September 1998)*, 3–8.

Moeschler, J. (2002). Economy and pragmatic optimality: the case of directional inferences. *GG@G, 3*, 1–21.

Nagatsuji, S. (2012). SHITE connector and dictionary form connector: about factors of pragmatic differences. In A. Yoshimura, A. Suga, & N. Yamamoto (Eds.), *Observing linguistic phenomena* (pp. 335–346). Tokyo: Eihousha.

O'Brien, S. (2002). Teaching Post-editing: A Proposal for Course Content. *Proceedings of the 6th EAMT Workshop on "Teaching Machine Translation". EAMT/BCS, UMIST*, (pp. 99–106). Manchester.

O'Brien, S. (2005). Methodologies for measuring the correlations between post-editing effort and machine translatability. *Machine Translation 19*, 37–58.

O'Brien, S. (2006). Pauses as indicators of cognitive effort in post-editing machine translation output. *Across Languages and Cultures, 7*(1), 1–21.

Senez, D. (1998). The machine translation help desk and the post-editing service. *Terminologie et Traduction, 1*, 289–295.

Shibatani, M. (2009). Japanese. In B. Comrie (Ed.), *The world's major languages* (2nd ed., pp. 741–764). Oxford: Routledge.

Somers, H. (1997). A practical approach to using machine translation software: "Post-editing the source text". *The Translator,3*, 193–212.

Sperber, D., & Wilson, D. (1986/1995). *Relevance: communication and cognition* (second ed.). Oxford: Blackwell.

Szpak, K. S. (2012). *Mapeando a busca por semelhança interpretativa com auxílio de rastreamento ocular: uma abordagem processual sobre o papel das codificações conceituais e procedimentais na construção de significado em textos traduzidos.* [Mapping the search for interpretive resemblance by eye tracking: a procedural approach to the role of conceptual and procedural encodings in the construction of meaning in translation texts]. M.A. thesis.

TAUS. (2014). *Machine translation post-editing guidelines.* Retreived from https://evaluation.taus.net/resources/guidelines/post-editing/machine-translation-post-editing-guidelines

Underwood, N., & Jongejan, B. (2001). Translatability checker: A tool to help decide whether to use MT. In *Processing of MT Summit VIII: Machine Translation in the Information Age*, Santiago de Compostela, Spain, 363–368.

Vasconcellos, M. (1986a). Functional consideration in the postediting of machine-translated output. (W. P. Lehmann, Ed.) *Computers and Translation, 11*(1), 21–38.

Vasconcellos, M. (1986b). Post-editing on-screen: Machine translation from Spanish into English. *Proceedings of Translating and the Computer, 8* (pp. 133–146). London: Aslib.

Wilson, D. (2011). The conceptual-procedural distinction: Past, present, and future. In V. Escandell-Vidal, M. Leonett, & A. Ahern. *Procedural meaning: Problems and perspectives* (pp. 3–28). London: Emerald Group Publishing Limited.

Wilson, D., & Sperber, D. (1993). Linguistic form and relevance. *Lingua, 90*, 1–25.