Translator Post-Editing and Subject-Matter Expert Revision versus Subject-Matter Expert Post-Editing and Translator Revision

Özlem Temizöz
Kocaeli University, Universidad Rovira i Virgili
Kocaeli, TURKEY
ozlemtemizoz@gmail.com

ABSTRACT

The present paper reports on a study which compares the speed and end-product quality in the following workflows: (1) post-editing the machine translation (MT) output of a technical text by subject-matter experts (engineers) and subsequent revision by professional translators, and (2) post-editing the MT output of a technical text by professional translators and subsequent revision by subject-matter experts. The findings suggest that the quality of the end-product is ultimately not only the result of the reviser's abilities and expertise/experience in the subject-matter but it is also affected by the quality of the post-editing to be revised. The results also show that, for technical texts, there is no significant difference between engineer post-editing + translator revision and translator post-editing + engineer revision with regard to speed. As for quality, revision brings about a quality improvement in both workflows. However, when the recurrent errors are penalized, the added value brought about by engineer-revision of translator-post-editing is significant, while it is not very significant when the recurrent errors are not penalized. Finally, for both groups, the quality of revision significantly differs when the recurring errors are penalized and not penalized.

Keywords: translation workflow, subject-matter experts in the translation/post-editing process, recurrent errors, revision, translation speed, translation quality.

1. Introduction

Rapid developments in technology, widespread use of the Internet and globalization have increased the volume of material to be translated. The use of technology (translation memories and machine translation systems) is one way of increasing efficiency. Incorporation of people who are not professional translators into the process is another method which has been employed to meet the global demand. The non-professional translators include volunteers who are members of social networks and NGOs and subject-matter experts.

In addition to the use of technology, producing higher-quality translations in shorter periods of time requires specialization in at least a few fields. However, not many translators can choose one or two areas to be specialized in. Even if this could be possible for some, the need to benefit from subject-matter experts' specialization still maintains its importance. This has made subject-matter experts and/or cooperation with them crucial in the translation/post-editing and revision processes.

With the incorporation of subject-matter experts in the translation/post-editing processes and in the context of collaboration between subject-matter experts and professional translators, we observe that the first draft of the translation/post-edited machine translation output is either created by the subject-matter experts and revised by the professional translators or the professional translators create the first draft and the output is revised by the subject matter-experts.

We conducted the present study to understand whether and/or how the speed and end-product quality differ in the following workflows:

1. Post-editing the machine translation (MT) output of a technical text by subject-matter experts and subsequent revision by professional translators, and

2. Post-editing the MT output of a technical text by professional translators and subsequent revision by subject-matter experts.

To this end, we compared the following parameters in each workflow:

1. processing speed of post-editing the MT outputs and their revision (based on words processed per minute, w/pm),

2. quality of the post-edited MT outputs and their revised versions (based on LISA QA Model 3.1.).

At this point, it is necessary to note that the present study aims to investigate only technical texts, and other text types are beyond the scope of the study.

2. The Role of Subject-Matter Experts in the Translation/Post-editing Process

Subject-matter expert – professional translator collaboration can be observed in the translation and localization industry. SDL Language Services mentions that they ensure quality and consistency by employing "more native-speaking subject-matter experts than any other company, meaning excellent standards of language and cultural awareness." Another language service provider SimulTrans mentions cooperation between linguists and subject-matter experts in the revision of translations as a significant component of their internal quality checks during the conduct of a project.

In an article published on the web-site of Tcworld, a magazine for international information management, Muegge (2011) emphasizes the requirement to "employ subject-matter experts to check the suitability of those translated terms on the client side." Language Scientific, a US-based translation and localization company, goes further and asks its employees to be both translators and subject-matter experts. HISPAWORDS, a language service provider in the pharmaceutical industry, notes that medical texts are "professionally translated by a specialized linguist and then proofread by an expert in the field, followed by a Quality Assurance process that ensures top accuracy, consistency and quality."

The language service provider McElroy Translation explains its three-step quality assurance (QA) process to provide its clients with the highest-quality output for their technical translations. They note that "all translations are sent to a highly qualified native-speaking subject-matter expert for the initial translation.

In terminology management, Le Néal (2001: 651) points to the key role played by subject-matter experts "to check the accuracy of the technical information contained in the entries and to comment on usage or on any terms" proposed by the bilingual terminologist in creating terminology.

Although we are aware that the quality standards and the quality assurance processes declared by the translation service providers on media might describe the ideal or optimum conditions rather than reflecting the real quality processes implemented, it is obvious that the involvement of subject-matter experts in the translation/post-editing processes is regarded as a significant component of quality assurance.

3. Methodology

3.1. Setting

The study is composed of two parts, each part involving two sets of experiments:

Part 1: Post-editing MT output of a technical text by:

1. subject-matter experts (engineers), and
2. professional translators.

Part 2: Revision of the post-edited MT outputs

1. Revision by professional translators of the texts post-edited by subject-matter experts, and
2. Revision by the subject-matter experts of the texts post-edited by professional translators.

3.2. Subjects

The subject-matter experts were graduates of various engineering departments and they had been working at various international automotive companies in Turkey for at least three years. They, therefore, hold the technical knowledge, expertise and experience in engineering. Engineers have Turkish as their mother tongue, and they are proficient in English. However, they received no training in translation. Owing to the international composition of their companies, they have to work in a multilingual atmosphere which makes translation an indispensable and a natural component of their work. Unlike professional translators, the type of translation work engineers engage in is rather in the form of dealing with (understanding, using part of it in verbal or written presentations, transferring it to or sharing with another stakeholder, etc.) various texts (e-mails, reports, informative documents, etc.) that usually contain technical information.

The professional translators were freelancers with translation experience of at least three years. As with the engineers, Turkish is the translators' mother tongue and English is the primary foreign language. They do not have any formal education or experience in engineering. The translators in the sample usually translate texts on social sciences and education. Some of them translate literary, academic and legal texts as well. One of them translates medical texts in addition to literary texts. A deliberate decision was made not to assign translators who were specialized in the translation of technical and/or engineering texts because the aim was to find out how the performance of translators who did not hold expertise in engineering would differ from that of subject-matter experts who have expertise in engineering.

Thus, the first significant distinction between the two groups is that the subject-matter experts are specialized in engineering, whereas the translators are not. The second difference is that translators work full-time in the translation market and principally make a living from translation, whereas engineers deal with translation as a component of their daily work.

3.3. The Source Text

The source text was a technical text in English, and it contained 587 words. It was taken from the International Dismantling Information System (IDIS) which contains technical instructions on dismantling end-of-life vehicles. The source text in English was pre-translated with Google Translate from English into Turkish. After pre-translation with Google Translate, it

turned into a 482-word Turkish text which was used as the test text in the study.

3.4. Method

A pilot study was carried out with two subject-matter experts and two professional translators in order to test the methodology and detect possible flaws in the design. After analyzing the results of the pilot study, the main study was conducted with ten engineers and ten professional translators.

3.4.1. Part 1: Post-editing Experiment

In Part 1 of the main study, we had ten subject-matter experts (engineers) and ten professional translators post-edit the same 482-word technical text in Turkish pre-translated from English with Google Translate. The two groups worked separately on the same MT output. The post-editors were asked to post-edit the MT output (provided as a Word.doc) using their own computers on which the screen recording software BB Flashback had been set up before. The participants were asked to complete the task in one-go without being interrupted. The subjects would have access to the source text, the Internet and online dictionaries during the post-editing task; however, they were not allowed to use any translation memory. They were asked to do the post-editing by performing the changes they wanted to introduce on the MT output provided for them instead of creating a separate target text. In order to neutralize as many variables as possible, we imposed a time restriction. Each subject was given forty minutes to post-edit the MT output. However, in case they might not finish the task in forty minutes, they were set free to take any extra time.

The speed of each post-editing performance was captured with the screen recordings. In Part 1, the aim was mainly to identify the quality differences between the MT output of a technical text post-edited by subject-matter experts and professional translators. The results of Part 1 (comparison of the quality of a technical text post-edited by engineers and professional translators) were discussed in Temizöz 2016a. In this paper, we will discuss the results of Part 2 of the study as well as presenting the data regarding the post-editing speed.

3.4.2. Part 2: Revision Experiment

In Part 2, we asked ten professional translators to revise ten subject-matter expert post-editings collected in Part 1. In addition, ten subject-matter experts were asked to revise ten professional translator post-editings. Both group of subjects performed the tasks separately, yet, under the same conditions. The revision tasks were carried out on the subjects' own computers at their usual work places, and they were recorded using BB Flashback. The subjects were allowed to use the Internet during the task; however, they were not allowed to use any translation memory. They were asked to do the revision by performing the changes they wanted to introduce on the post-edited text provided for them instead of creating a separate target text. The subjects were also provided with the source text as well for reference.

The subject-matter experts and the professional translators doing the revision tasks were the same with those who carried out the post-editing tasks. In other words, subject-matter expert post-editors and professional translators post-edited the same MT output. Then, both groups of revisers that were composed of the same participants in the post-editing task revised each other's post-edited texts. This might result in their learning of the text that they have already post-edited. However, since revising a text in which the reviser has already had experience is very natural and common in real life, this learning factor does not pose a problem for our research. Nevertheless, in an attempt to avoid the learning effect, we conducted the revision experiment two months after the post-editing experiment.

3.5. Quality Analysis Procedure

The quality analysis procedure of the post-edited texts and their revised versions was the same. Both the post-edited target texts and the revised versions were compared with the reference translation of the test text from English into Turkish. This comparison was made by the author, and the reference translation was performed by a professional translator (with a PhD in translation and ten years of experience in the profession) and a mechanical engineer (with a TOEIC score of 900 out of 990 and ten years of experience in engineering at an international automotive company) in cooperation. The translator and the engineer performing the reference translation did not take part in the experiment.

Since each post-editor/translator might translate the same text differently (even the same translator may translate the same text in slightly different ways at different times), when determining errors in the post-edited texts, we did not look for exactly the same words or expressions occurring in the reference translation. Since the test text was a technical text and not open to interpretations, we did not have to deal with post-editings/revisions that were very different from the reference translation.

LISA QA Model 3.1 was used as a tool for measuring quality. Both groups' error points and quality percentages obtained from LISA QA Model 3.1 interface were compared. For post-editing, the minimum acceptable level of quality was set as 75 per cent, while it was set as 85 per cent for revision. The reason for this 10-per cent increase in the quality threshold is that the revised texts are the final versions of the post-editings; thus, their quality requirement is higher.

3.5.1. Quality of the Raw Machine Translation (MT) Output

The quality of the raw (unedited) machine translation output is an important factor that affects the quality of the post-edited text.

The MT output used in this study received a BLEU score of 0.3465 (Papineni et al. 2002). This score indicates an acceptable level of quality if we consider that "on a test corpus of about 500 sentences, a human translator scored 0.3468 against four references and scored 0.2571 against two references" (Papineni et al., 2002, p. 315). However, it contained errors and required post-editing. Because of the relative deficiency of the MT database in Turkish, it was fed with the translations of a few texts that belong to the same genre as the experiment text. These translations were performed by a professional translator with technical translation experience by using Google Translator Toolkit (GTT) and with access to the Internet.

3.6. Measuring Processing Speed

Speed was measured with two types of data: total processing speed captured with screen recording software and the number of words processed per minute (w/pm). Four measures were taken: (1) the processing speed of subject-matter experts' post-editing MT outputs, (2) the processing speed of professional translators' post-editing MT outputs, (3) the processing speed of revision of subject-matter experts' post-editings by professional translators, (4) the processing speed of revision of professional translators' post-editings by subject-matter experts.

For items 3 and 4, data were recorded for the revision tasks only (only translators' revisions or only engineers' revisions) as well as for the whole task involving post-editing and revision (engineers' post-editing + translators' revision and translators' post-editing + engineers' revision).

When calculating the w/pm for the post-editing task, we used the word-count of the raw MT output (482) for all the post-editors. However, for revision, the number of words each reviser processes depends on the number of words produced by the post-editor. That is, the word-count of the raw MT output is changed during post-editing with additions and deletions made by the post-editor. Thus, for the calculation of w/pm for each revised text, we used the different word-

counts produced by each post-editor. In order to calculate the w/pm for post-editing + revision tasks, we considered both the word count of the raw MT output (for w/pm of post-editing) and that of the post-edited text (for w/pm of revision).

4. Findings on Revision (Part 2)

4.1. Speed in Revision

Table 1 presents the speed of translator post-editing in comparison with its revision by engineers, the speed of engineer post-editing in comparison with its revision by translators as well as the comparison of the two workflows' (translator post-editing + engineer revision and engineer post-editing + translator revision) total task time comprising post-editing and revision. Both groups completed the revision task in almost half as much the time they completed the post-editing task. On average, translators spent 62.70 minutes for post-editing and 27.40 minutes for revision (Table 1). On the other hand, engineers spent 55.66 minutes for post-editing and 33.31 minutes for revision. Unlike the mean duration of post-editing, in which engineers worked faster than translators (55.66 minutes against 62.70 minutes), the mean duration of the engineers' revision is longer than that of the translators' (33.31 minutes against 27.40 minutes). In revision, engineers processed, on average, 18.6 w/pm, while translators processed 29.2 w/pm. Although the mean values indicate that the translators revise faster than the engineers, according to the t-test, there is no significant difference between translators and engineers with regard to speed in revising one another's post-editing. The p-value for the comparison of both groups' revision duration is 0.4640 and it is 0.2727 for their w/pm (Table 1).

In addition to comparing the speed of only the revision tasks, we compared the speed of post-editing + revision processes. The mean duration of engineer post-editing + translator revision is thirteen minutes shorter than translator post-editing + engineer revision. On average, the former process took 83.06 minutes, while the latter took 96.01 minutes. The mean value of w/pm for engineer post-editing + translator revision is 13.2, and this figure is 10.7 for translator post-editing + engineer revision (Table 1). Although the mean values for the total task time comprising post-editing + revision indicate that engineer post-editing + translator revision is faster than translator post-editing + engineer revision, the t-tests do not show any significant difference between both processes with regard to speed (p-value for post-editing + revision task time is 0.3021 and it is 0.1611 for w/pm).

Table 1. Speed in Revision - Translators and Engineers

Translator Post-editing			Engineer Revision of Translator Post-editing				
Subjects	PE Task Time incl. self-revision (minute)	Word processed per minute	Subjects	REV Task Time (minute)	Word processed per minute	Translator PE+Engineer Revision Task Time (minute)	Translator PE+Engineer Revision w/pm
Tr 1	44.98	10.7	En 1	24.73	20.5	69.71	14.2
Tr 2	67.53	7.1	En 2	42.41	12.1	109.94	9.1
Tr 3	45.36	10.6	En 3	54.73	8.7	100.10	9.6
Tr 4	60.16	8.0	En 4	12.70	39.1	72.86	13.4
Tr 5	107.75	4.5	En 5	14.50	32.2	122.25	7.8
Tr 6	38.73	12.4	En 6	22.66	21.5	61.40	15.8
Tr 7	68.75	7.0	En 7	28.93	16.8	97.68	9.9
Tr 8	62.66	7.6	En 8	58.91	8.2	121.58	7.9
Tr 9	60.03	8.0	En 9	38.48	12.9	98.51	10.0
Tr 10	71.08	6.7	En 10	35.03	13.7	106.11	9.1
Mean	62.70	8.3		33.31	18.6	96.01	10.7
Engineer Post-editing			Translator Revision of Engineer Post-editing				
Subjects	PE Task Time incl. self-revision (minute)	Word processed per minute	Subjects	REV Task Time (minute)	Word processed per minute	Engineer PE+Translator Revision Task Time (minute)	Engineer PE+Translator Revision w/pm
En 1	52.96	9.1	Tr 1	26.21	18.3	79.18	12.2
En 2	47.45	10.1	Tr 2	18.91	26.1	66.36	14.7
En 3	46.00	10.4	Tr 3	24.93	20.0	70.30	13.9
En 4	65.61	7.3	Tr 4	32.53	14.7	98.15	9.8
En 5	60.96	7.9	Tr 5	40.26	11.6	101.23	9.4
En 6	44.51	10.8	Tr 6	16.50	29.9	61.01	16.0
En 7	61.70	7.8	Tr 7	18.40	27.9	80.10	12.4
En 8	84.30	5.7	Tr 8	75.28	6.8	159.58	6.3
En 9	35.40	13.6	Tr 9	04.55	105.0	39.95	24.0
En 10	57.73	8.3	Tr 10	16.40	31.5	74.13	13.5
Mean	55.66	9.1		27.40	29.2	83.06	13.2
p-value	0.3598	0.4195		0.4640	0.2727	0.3021	0.1611

4.2. Quality in Revision

We evaluated the quality of revision in two different ways. First, we compared the quality of revisions performed by professional translators and subject-matter experts (engineers). Second, we compared the quality of the post-edited texts and their final revised versions in order to determine how the subject-matter knowledge of translators/engineers as revisers affects the quality of engineers'/translators' post-editing.

The study also revealed that post-editing quality was significantly affected by counting or not counting the recurrent errors in the text (Temizöz 2016b). The test text was a technical text, and it was not a deliberate decision to choose one which contained recurrent terms. However, being repetitive is among the characteristics of technical texts. Therefore, rather than changing the text, we decided to approach quality in two ways: by counting the recurring errors each time they occur, and by disregarding the recurrent versions of the same error and count them only once (Temizöz 2016b). As a result, as is the case in post-editing, due to the recurrent nature of the errors, the number of errors, and thus the quality of revision as measured by LISA QA 3.1 might change when we penalize or do not penalize the recurring errors.

Therefore, we decided to look at the quality results by penalizing and not penalizing the recurrent errors.

4.2.1. Quality in Revision - Recurring Errors Penalized

The revised texts were compared with the reference translation of the source text with a view to detecting errors. The errors were classified in line with LISA error categories. As it had been the case in post-editing (Part 1), there were no errors in "style" and "country" categories, so they were removed from the error analysis.

In Part 1, we found that, for this particular task (technical translation), translators' and engineers' post-editing quality was similar as far as the categories of mistranslation, accuracy, and consistency were concerned. Engineers performed significantly better than translators only in the terminology category (Temizöz 2016a). However, the difference between both groups in terminology category affected the overall quality of the post-edited texts as measured by LISA QA Model. Thus, LISA grid showed that engineers' post-editing of a technical text was of significantly higher quality than the translators ($p = 0.0339$, Table 47) when the outlier participants were removed. When the outliers were not removed, the p-value was 0.0636, Table 5), which was also near the significance level (Temizöz 2016a).

In order to analyze revision quality, first, we compared the revision processes performed by translators and engineers. To this end, we looked at the total error points (calculated by LISA QA Model based on the errors entered) and the quality percentages of translators' and engineers' revision (Table 2). According to the mean values, the engineers' total error points is 94 and translators' is 101. On average, the quality percentage of engineers' revision is 81.05, while that of translators' is 79.68. A total of six out of ten engineers passed the quality threshold set for revision (85 per cent), while a total of five out of ten translators passed this threshold. Although the mean values indicate that engineer revision is of slightly

higher quality than translator revision, the t-test shows no significant difference (p = 0.8689 for the total error points and 0.8685 for the quality percentages, Table 2). However, there was a significant difference between both groups' post-editing qualities in that the engineers' post-editing was of significantly higher quality than translators'. This decrease in the quality difference between translators and engineers in revision can be explained by the fact that the higher quality of engineer post-editing affected the quality of translator revision positively, while the relatively lower quality of translator post-editing affected the quality of engineer revision negatively. This reveals that revision quality is dependent not only on the qualifications of the reviser but also on the quality of the text to be revised.

Table 2. Quality in Revision - Translators and Engineers - Recurring Errors Penalized

Subjects	Total Error Points	Revision Quality %
Engineers' Mean	94	81.05
Translators' Mean	101	79.68
p-value	0.8689	0.8685

Cooperation between professional translators and subject-matter experts is expected to bring higher-quality texts. However, as is seen above, not all types of cooperation bring higher quality. The quality of the end-products revised by one group is affected by the strong and weak aspects of the post-editing by the other group. However, we can test which workflow brings more added value to the quality of the end product: translator post-editing + engineer revision or engineer post-editing + translator revision. By "added value", we mean the quality increase brought about by revision. Thus, the third and the most important aspect of Part 2 is to determine the effect of the reviser's subject-matter knowledge on the quality of the post-edited text. Table 3 compares the total error points and the quality percentages of translator post-editing with engineer revision. An overview of the data shows that – except for Engineer 9 – engineer revision increased the quality of translator post-editing. With engineer revision, the mean total error points in translator post-editing decreased from 187 to 94, which is a significant difference (p = 0.0205). The quality is increased by 19.78 per cent (from 61.27 to 81.05). Thus, engineer revision significantly increased the quality of translator post-editing (p = 0.0164, Table 3), indicating that the subject-matter knowledge of the reviser is more important in determining the end-product quality than is that of the post-editor.

Table 3. Translator Post-editing and Engineer Revision - Quality Compared - Recurring Errors Penalized

Translator Post-editing		
Subjects	Total Error Points	Post-editing Quality %
Mean	187	61.27
Engineer Revision of Translator Post-editing		
Subjects	Total Error Points	Revision Quality %
Mean	94	81.05
p-value	0.0205	0.0164

We also compared the quality of engineer post-editing with its revision by translators. An overview of the data shows that – except for Translator 2 – translator revision increased the quality of engineer post-editing. With translator revision, the mean total error points in engineer post-editing decreased from 118 to 101. There is a total of 17 points decrease, which does not correspond to a significant difference (p = 0.6626, Table 4). The quality percentage is increased by only 4.10 per cent (from 75.58 to 79.68). As a result, although the mean values show that translator revision increased the quality of engineer post-editing, it is not possible to talk about a significant increase (p = 0.6100, Table 4). This result affirms our finding above that the subject-matter knowledge of the reviser is more important in determining the end-product quality than is that of the post-editor.

Table 4. Engineer PE + Translator Revision - Quality Compared - Recurring Errors Penalized

Engineer Post-editing		
Subjects	Total Error Points	Post-editing Quality %
Mean	118	75.58
Translator Revision of Engineer Post-editing		
Subjects	Total Error Points	Revision Quality %
Mean	101	79.68
p-value	0.6626	0.6100

Table 5 compares the quality difference between one group's post-editing and the other group's revision. Most of the data in Table 5 were presented in the previous tables in this section. However, unlike the previous ones, Table 5 compares the quality differences brought about by revision in the two different processes: translator post-editing + engineer revision and engineer post-editing + translator revision. Engineer revision decreased the total error points of translator post-editing by 93 points, and it increased the quality of translator post-editing by 19.78 per cent. On the other hand, translator revision decreased the total error points of engineer post-editing by 17 points, and it increased the quality of engineer post-editing by 4.10 per cent. Consequently, when the recurring errors are penalized, the quality improvement brought about by engineer-revision of translator-post-editing is significantly higher than the quality improvement brought about by translator-revision of engineer-post-editing (p = 0.0073 for the total error points and p = 0.0061 for the quality percentage, Table 5).

Table 5. Quality Difference Between Post-editing and Its Revision - Recurring Errors Penalized

Translator Post-editing			Engineer Revision of Translator Post-editing				
Subjects	Total Error Points	PE Quality %	Subjects	Total Error Points	Revision Quality	Difference btw. REV and PE	Difference btw. REV and PE

					%	Error Pts.	Quality Percentage %
Translators ' Mean	187	61.27	Engineers' Mean	94	81.05	-93	19.78
Engineer Post-editing			Translator Revision of Engineer Post-editing				
Engineers' Mean	118	75.58	Translators' Mean	101	79.68	-17	4.10
p-value	0.0636		p-value	0.8689	0.8685	0.0073	0.0061

4.2.2. Quality in Revision - Recurring Errors Not Penalized

Previous studies show that the quality of the translated/post-edited text changes significantly depending on, during the quality analysis, whether the recurring errors are counted each time they occur in the text or only once, disregarding the recurrent versions of the same error (O'Brien 2012, Temizöz 2016b). Therefore, we repeated the analysis of revision quality by also not penalizing the recurring errors to see if the results change under this condition.

First, we compared the revision processes performed by translators and engineers. To this end, we looked at the total error points (calculated by the LISA QA Model based on the errors entered) and the quality percentages of translators' and engineers' revision (Table 6). According to the mean values, the engineers' total error points is 32 and translators' is 31. On average, the quality percentage of engineers' revision is 93.55, while that of translators' is 93.76. All of the translators passed the quality threshold set for revision (85 per cent), while a total of nine out of ten engineers passed this threshold. The engineer who failed had a quality percentage of 84.25, which is very close to the threshold (Table 11). Under the penalized condition, however, a total of five out of ten translators and six out of ten engineers passed the quality test (Table 10 and Table 11). Nevertheless, the difference between the translators' and engineers' revision, as is the case under the penalized condition (see Table 2 for $p = 0.8689$ for the total error points and $p = 0.8685$ for the quality percentage), the t-test shows no significant difference between the translators' and engineers' revision qualities under the not penalized condition ($p = 0.9162$ for the total error points and 0.9134 for the quality percentages, Table 6).

Table 6. Quality in Revision - Translators and Engineers - Recurring Errors Not Penalized

Recurring Errors Not Penalized		
Subjects	Total Error Points	Revision Quality %
Engineers' Mean	32	93.55
Translators' Mean	31	93.76
p-value	0.9162	0.9134

We also compared the end-products of the two workflows (translator post-editing + engineer revision or engineer post-editing + translator revision) in terms of the added value when the recurrent errors were not penalized. By "added value" we mean the quality increase brought about by revision. By means of this analysis, we aim to determine the difference between post-editing and revision qualities

under the not penalized condition, which will then be compared with the penalized condition.

Table 7 compares the total error points and the quality percentages of translator post-editing with engineer revision. An overview of the data shows that engineer revision increased the quality of translator post-editing. With engineer revision, the mean total error points in translator post-editing decreases from 52 to 32; however, the difference is not very significant (p = 0.0880). The quality percentage is increased by 4.25 per cent (from 89.30 to 93.55), which is not very significant either (p = 0.0751). Thus, unlike the penalized condition where the p-value for the difference between the quality percentages of translator post-editing and engineer revision is 0.0164 (Table 3), engineer revision did not significantly increase the quality of translator post-editing when the recurrent errors were not penalized (p = 0.0751, Table 7). On the other hand, since it is close to the threshold 0.05, a p-value of 0.0751 does not represent a very low level of significance. We can explain the reason for this difference between the penalized and not penalized conditions as follows: The engineers make fewer recurrent terminology errors than translators; thus, the increase in the engineers' quality score brought about by not penalizing the recurrent errors is low. On the other hand, the translators make more recurrent terminology errors than engineers; thus, the increase in the translators' quality score brought about by not penalizing the recurrent errors is high. Consequently, when the recurrent errors are not penalized, the difference between engineer revision and translator post-editing diminishes as compared to the penalized condition.

Table 7. Translator Post-editing and Engineer Revision - Quality Compared - Recurring Errors Not Penalized

Translator Post-editing		
Subjects	Total Error Points	Post-editing Quality %
Translators' Mean	52	89.30
Engineer Revision of Translator Post-editing		
Subjects	Total Error Points	Revision Quality %
Engineers' Mean	32	93.55
p-value	0.0880	0.0751

As for the comparison of the quality of engineer post-editing with its revision by translators, our data show that, under the not penalized condition, translator revision increased the quality of engineer post-editing, too. With translator revision, the mean total error points in engineer post-editing decreased from 45 to 31. There is a total of 14 points decrease, which does not correspond to a significant difference (p = 0.0827, Table 8). The quality percentage is increased by 3.03 per cent (from 90.73 to 93.76) which is not very significant, either (p = 0.0685). However, this is a rather significant figure, at least as compared to the penalized condition where the p-value of the difference between the quality percentages of engineer post-editing and translator revision is 0.6100 (Table 4). As stated above, this difference between the penalized and not penalized conditions is caused by the higher number recurrent terminology errors made by translators in comparison with engineers. All in all, it is important to note that the level of significance of the difference between engineer post-editing and translator revision increases when we do not penalize the recurrent errors.

Table 8. Engineer Post-editing and Translator Revision - Quality Compared - Recurring Errors Not Penalized

Engineer Post-editing		
Subjects	Total Error Points	Post-editing Quality %
Engineers' Mean	45	90.73
Translator Revision of Engineer Post-editing		
Subjects	Total Error Points	Revision Quality %
Translators' Mean	31	93.76
p-value	0.0827	0.0685

Table 9 compares the quality difference between one group's post-editing and the other group's revision when the recurrent errors are not penalized. Most of the data in Table 9 were presented in the previous tables in this section. However, unlike the previous tables, Table 9 compares the quality differences brought about by revision in two different workflows: translator post-editing + engineer revision and engineer post-editing + translator revision. In translator post-editing + engineer revision, on average, engineer revision decreased the total error points of translator post-editing by 20 points, and it increased the quality of translator post-editing by 4.25 per cent (Table 9). Under the recurring errors penalized condition, these figures were 93 and 19.78, respectively (Table 5). On the other hand, in engineer post-editing + translator revision, translator revision decreased the total error points of engineer post-editing by 14 points, and it increased the quality of engineer post-editing by 3.03 per cent (Table 9). Under the recurring errors penalized condition, these figures were 17 and 4.10 respectively (Table 5).

Under the recurring errors not penalized condition, there is no significant difference between the quality improvement brought about by engineer-revision of translator-post-editing and translator revision of engineer post-editing (the p-value for the total error points is 0.1458 and for the quality percentage 0.1219, Table 9). However, under the penalized condition, the quality improvement brought about by engineer-revision of translator-post-editing was significantly higher than the quality improvement brought about by translator-revision of engineer-post-editing ($p = 0.0073$ for the total error points and $p = 0.0061$ for the quality percentage, Table 5).

Table 9. Quality Difference Between Post-editing and Its Revision - Recurring Errors Not Penalized

Translator Post-editing			Engineer Revision of Translator Post-editing				
Subjects	Total Error Points	Post-editing Quality %	Subjects	Total Error Points	Revision Quality %	Difference btw. REV and PE Error Pts.	Difference btw. REV and PE Quality Percentage %
Translators' Mean	52	89.30	Mean	32	93.55	-20	4.25
Engineer Post-editing			Translator Revision of Engineer Post-editing				

Engineers' Mean	45	90.73	Mean	31	93.76	-14	3.03
p-value	0.4711	0.4708	p-value	0.9162	0.9134	0.1458	0.1219

In order to gain a clearer view of how each group's data compare under the penalized and not penalized conditions, we compared the total error points and the quality percentages of each group under both conditions. Table 10 compares the data on the translators' revision of engineer post-editing under the penalized and not penalized conditions. It is interesting to note that there is a significant difference between the penalized and not penalized conditions with regard to the total error points (p = 0.0257) and the quality percentages (p = 0.0265, Table 10). However, penalized and not penalized conditions do not differ significantly as far as the added value brought about by translator-revision of engineer-post-editing is concerned (p = 0.5909 for the difference between revision and post-editing total error points and 0.3841 for the difference between revision and post-editing quality percentages, Table 10).

Table 10. Translators' Revision Quality - Penalized and Not Penalized Conditions

Recurring Errors Penalized									
Engineer Post-editing				Translator Revision of Engineer Post-editing					
Subjects	Total Error Points	Post-editing Quality %	Result	Subjects	Total Error Points	Revision Quality %	Result	Difference btw. REV and PE Error Pts.	Difference btw. REV and PE Quality Percentage %
En 1	42	91.29	Pass	Tr 1	32	93.31	Pass	-10	2.02
En 2	152	68.46	Fail	Tr 2	162	67.73	Fail	10	-0.73
En 3	225	53.32	Fail	Tr 3	219	56.55	Fail	-6	3.23
En 4	57	88.17	Pass	Tr 4	2	99.59	Pass	-55	11.42
En 5	80	83.40	Pass	Tr 5	76	83.80	Fail	-4	0.40
En 6	235	51.24	Fail	Tr 6	210	57.75	Fail	-25	6.51
En 7	55	88.59	Pass	Tr 7	40	92.31	Pass	-15	3.72
En 8	65	86.51	Pass	Tr 8	34	93.54	Pass	-31	7.03
En 9	228	52.70	Fail	Tr 9	210	56.16	Fail	-18	3.46
En 10	38	92.12	Pass	Tr 10	21	96.09	Pass	-17	3.97
Mean	118	75.58		Mean	101	79.68		-17	4.10
Recurring Errors Not Penalized									
Engineer Post-editing				Translator Revision of Engineer Post-editing					
Subjects	Total Error Points	Post-editing Quality %	Result	Subjects	Total Error Points	Revision Quality %	Result	Difference btw. REV and PE Error Pts.	Difference btw. REV and PE Quality Percentage %
En 1	34	92.95	Pass	Tr 1	21	95.61	Pass	-13	2.66
En 2	40	91.70	Pass	Tr 2	32	93.63	Pass	-8	1.93
En 3	54	88.80	Pass	Tr 3	46	90.87	Pass	-8	2.07
En 4	32	93.36	Pass	Tr 4	2	99.59	Pass	-30	6.23
En 5	66	86.31	Pass	Tr 5	48	89.77	Pass	-18	3.46
En 6	71	85.27	Pass	Tr 6	53	89.34	Pass	-18	4.07
En 7	33	93.15	Pass	Tr 7	14	97.31	Pass	-19	4.16
En 8	29	93.98	Pass	Tr 8	22	95.82	Pass	-7	1.84
En 9	58	87.97	Pass	Tr 9	50	89.56	Pass	-8	1.59
En 10	30	93.78	Pass	Tr 10	21	96.09	Pass	-9	2.31
Mean	45	90.73		Mean	31	93.76		-14	3.03
p-value	0.0140	0.0140			0.0257	0.0265		0.5909	0.3841

Table 11 compares the data on the engineers' revision of translator post-editing under the penalized and not penalized conditions. As is the translators' revision quality, engineers' revision quality differs significantly under the penalized and not penalized conditions ($p = 0.0497$ for the total error points and $p = 0.0508$ for the quality percentage, Table 11). Unlike the translators' revision, engineers' revision quality also differ significantly under the penalized and not penalized conditions with regard to the added value brought about by engineer-revision of translator-post-editing ($p = 0.0082$ for the difference between revision and post-editing total error points and 0.0058 for the difference between revision and post-editing quality percentages, Table 11).

Consequently, our findings indicate that penalizing or not penalizing recurrent errors affects the revision quality significantly.

Table 11. Engineers' Revision Quality - Penalized and Not Penalized Conditions

Recurring Errors Penalized									
Translator Post-editing				Engineer Revision of Translator Post-editing					
Subjects	Total Error Points	Post-editing Quality %	Result	Subjects	Total Error Points	Revision Quality %	Result	Difference btw. REV and PE Error Pts.	Difference btw. REV and PE Quality Percentage %
Tr 1	55	88.59	Pass	En 1	24	95.34	Pass	-31	6.75
Tr 2	180	62.66	Fail	En 2	171	66.54	Fail	-9	3.88
Tr 3	258	46.47	Fail	En 3	140	72.11	Fail	-118	25.64
Tr 4	80	83.40	Pass	En 4	6	98.78	Pass	-74	15.38
Tr 5	241	50.00	Fail	En 5	64	86.21	Pass	-177	36.21
Tr 6	284	41.08	Fail	En 6	256	47.65	Fail	-28	6.57
Tr 7	197	59.13	Fail	En 7	32	93.59	Pass	-165	34.46
Tr 8	178	63.07	Fail	En 8	21	96.09	Pass	-157	33.02
Tr 9	182	62.24	Fail	En 9	200	59.35	Fail	18	-2.89
Tr 10	212	56.02	Fail	En 10	25	94.80	Pass	-187	38.78
Mean	187	61.27		Mean	94	81.05		-93	19.78
Recurring Errors Not Penalized									
Translator Post-editing				Engineer Revision of Translator Post-editing					
Subjects	Total Error Points	Post-editing Quality %	Result	Subjects	Total Error Points	Revision Quality %	Result	Difference btw. REV and PE Error Pts.	Difference btw. REV and PE Quality Percentage %
Tr 1	23	95.23	Pass	En 1	17	96.70	Pass	-6	1.47
Tr 2	45	90.66	Pass	En 2	20	96.09	Pass	-25	5.43
Tr 3	85	82.37	Pass	En 3	68	86.45	Pass	-17	4.08
Tr 4	32	93.36	Pass	En 4	6	98.78	Pass	-26	5.42
Tr 5	81	83.20	Pass	En 5	39	91.59	Pass	-42	8.39
Tr 6	91	81.12	Pass	En 6	77	84.25	Fail	-14	3.13
Tr 7	39	91.91	Pass	En 7	18	96.39	Pass	-21	4.48
Tr 8	35	92.74	Pass	En 8	19	96.46	Pass	-16	3.72
Tr 9	53	89.00	Pass	En 9	40	91.87	Pass	-13	2.87
Tr 10	32	93.36	Pass	En 10	15	96.88	Pass	-17	3.52
Mean	52	89.30	Pass	Mean	32	93.55		-20	4.25
p-value	0.0001	0.0001			0.0497	0.0508		0.0082	0.0058

5. Conclusions

In the present study, the following workflows were compared in terms of speed and end-product quality:

1. Post-editing the MT output by subject-matter experts and subsequent revision by professional translators, and

2. Post-editing the MT output by professional translators and subsequent revision by subject-matter experts.

Speed was measured as words processed per minute in both the post-editing and revision tasks, and quality was measured by evaluating the revised end-products using the LISA QA Model 3.1.

The study yielded interesting results:

1. There is no significant difference between translators and engineers with regard to speed in revising one another's post-editing (the p-value for revision duration is 0.4640 and 0.2727 for their w/pm, Table 1).

2. There is no significant difference between engineer post-editing + translator revision and translator post-editing + engineer revision with regard to speed (the p-value for post-editing + revision task time is 0.3021 and it is 0.1611 for w/pm, Table 1).

As for output quality, when the recurrent errors are penalized:

1. Engineer revision significantly increased the quality of translator post-editing ($p = 0.0164$, Table 3), indicating that the subject-matter knowledge of the reviser is more important in determining the end-product quality than is that of the post-editor.

2. Translator revision increased the quality of engineer post-editing, yet it is not possible to talk about a significant increase ($p = 0.6100$, Table 4). This result affirms our finding above that the subject-matter knowledge of the reviser is more important in determining the end-product quality than is that of the post-editor.

3. Consequently, for technical texts, when the recurrent errors are penalized, the quality improvement brought about by engineer-revision of translator-post-editing is significantly higher than the quality improvement brought about by translator-revision of engineer-post-editing (the p-value for the total error points is 0.0073 and for the quality percentages 0.0061, Table 5).

When the recurrent errors are not penalized:

1. As under the penalized condition (Table 2), there is no significant difference between the translators' and engineers' revision qualities when the recurrent errors are not penalized (p = 0.9162 for the total error points and 0.9134 for the quality percentages, Table 6).

2. Engineer revision increased the quality of translator post-editing, yet unlike the penalized condition (p = 0.0164, Table 3), the difference is not significant (p = 0.0751, Table 7).

3. Translator revision increased the quality of engineer post-editing. Although the difference is not quite significant (p = 0.0685, Table 8), it is close to the threshold 0.05 as compared to the penalized condition, where the p-value for the difference between engineer post-editing and translator revision of engineer post-editing is 0.6100 (Table 4).

4. Unlike the penalized condition (the p-value for the total error points is 0.0073 and for the quality percentages 0.0061, Table 5), the quality improvement brought about by engineer-revision of translator-post-editing does not significantly differ from the quality improvement brought about by translator-revision of engineer-post-editing (the p-value for the total error points is 0.1458 and for the quality percentages 0.1219, Table 9).

5. The translators' revision quality significantly differs under the penalized and not penalized conditions (p = 0.0265, Table 10). However, as far as the added value brought about by translator-revision of engineer-post-editing is concerned, the two conditions (penalized and not penalized) do not differ significantly (p = 0.3841, Table 10).

6. The engineers' revision quality significantly differs under the penalized and not penalized conditions (p = 0.0508, Table 11). Unlike translator revision of engineer post-editing, the conditions where recurrent errors are penalized and not penalized also differ significantly as far as the added value brought about by engineer-revision of translator-post-editing is concerned (p = 0.0058, Table 11).

Consequently, the mean differences between post-editing and revision indicate that revision brings about a quality improvement in both workflows (translator post-editing + engineer revision and engineer post-editing + translator revision) under both conditions (penalized and not penalized) (Table 5 and Table 9). However, when the recurrent errors are penalized, the added value brought about by engineer-revision of translator-post-editing is significant (p = 0.0164, Table 3), while it is not very significant when the recurrent errors are not penalized (p = 0.0751, Table 7). As for the added value brought about by translator-revision of engineer-post-editing, we did not find a significant quality difference between the post-editings and revisions, either under the penalized (p = 0.6100, Table 4) or not penalized conditions (p = 0.0685, Table 8). However, it is important to note that the significance of the difference between translator post-editing and engineer revision decreased under not penalized condition (p = 0.0164, Table 3 compared to p = 0.0751, Table 7), while the significance of the difference between engineer post-editing and translator revision increased (p = 0.6100, Table 4 compared to p = 0.0685, Table 8). Moreover, under the penalized condition, the quality improvement brought about by engineer-revision of translator-post-editing is significantly higher than vice-versa (p = 0.0061, Table 5). When the recurrent

errors are not penalized, however, the difference between the quality improvement brought about by engineer revision and translator revision becomes insignificant ($p = 0.1219$, Table 9). Finally, for both groups, the quality of revision significantly differs when the recurring errors are penalized and not penalized (Table 10 and Table 11).

6. Final Conclusions and Recommendations for the Translation Industry

The findings of this research suggest that, for technical texts, the higher quality of engineer post-editing has a positive impact on the quality of translator revision and the relatively lower quality of translator post-editing has a negative impact on the quality of engineer revision. In other words, the final quality of the end-product is not only the result of the reviser's abilities and expertise/experience in the subject-matter but it is also affected by the quality of the post-editing to be revised. Thus, when trying to decide which workflow (translator post-editing + subject-matter expert revision or subject-matter expert post-editing + translator revision) produces higher-quality outputs, we need to consider this effect of the quality of the text to be revised on the final output quality. So translation companies should bear in mind that recruiting subject-matter experts for the final revision as a method of quality assurance may not always give the expected results. Rather, they should seek experience and expertise in the subject-matter both for the translator/post-editor and for the reviser if they want to ensure higher-quality outputs.

Our data also show that, for technical texts, the quality improvement (added value) brought about by engineer-revision of translator-post-editing is higher than the quality improvement brought about by translator-revision of engineer-post-editing. However, this added value is significant only when the recurrent errors are penalized, and it is not very significant when the recurrent errors are not penalized.

Penalizing and not penalizing recurrent errors are the methods employed in the translation industry for determining the output quality mostly without being questioned. This research shows that both the quality of the post-edited texts and their revised versions (either performed by professional translators or subject-matter experts) change significantly as a result of penalizing and not penalizing recurrent errors (see also O'Brien 2012, Temizöz 2016b). Thus, in order to obtain sound and consistent quality results, the translation industry should take this significant but rather ignored point into consideration when evaluating the quality of the texts post-edited/translated or revised.

7. References

BB FlashBack (2011) Screen Recorder Software, Blueberry Software Ltd. URL: http://www.bbsoftware.co.uk/bbflashback/home.aspx. (accessed 12 January 2011)

Google Translator Toolkit (2010) "Home page of Google Translator Toolkit". http://translate.google.com/toolkit. (accessed 20 November 2010).

HISPAWORDS (2011) "Home page of HISPAWORDS Language Service Provider". http://hispawords.com/life_sciences.html. (accessed 15 September 2011)

IDIS (2011) "Home page of the International Dismantling Information System". http://www.idis2.com. (accessed 6 June 2011)

Language Scientific (2011) "Home page of Language Scientific, Translation and

Localization Company". http://www.languagescientific.com/translation-quality-control-process.html. (accessed 16 September 2011)

Le Néal, Jocelyne (2001) "Preparing Multi-Volume Illustrated Terminological Dictionaries" in *Handbook of Terminology Management. Volume II: Application-Oriented Terminology Management*, S.E. Wright, and G. Budin (eds), Amsterdam and Philadelphia: John Benjamins: 645-665.

LISA (2010) "Home page of the Localization Industry Standards Association". http://www.lisa.org/LISA-QA-Model-3-1.124.0.html. (accessed 7 July 2010)

McElroy Translation (2011) "Home page of McElroy Translation".http://www.mcelroytranslation.com/processes/qualityassurance/workflowprocess. (accessed 11 September 2011)

Muegge, Uwe (2011) "Ten good reasons why you should validate your translated terminology". http://www.tcworld.info/tcworld/translation-and-localization/article/ten-good-reasons-why-you-should-validate-your-translated-terminology. (accessed 10 September 2011)

O'Brien, Sharon (2012) "Towards a Dynamic Quality Evaluation Model for Translation", Journal of Specialized Translation 17, no. 1: 1-24.

Papineni et al. (2002), "BLEU: A method for automatic evaluation of machine translation", In Proceedings of the 40th Annual Meeting of the Association for Computational Linguistics (ACL), Philadelphia: 311-318.

SDL (2011) "Home page of SDL Language Services". http://www.sdl.com/en/language-services/translation-solutions/translation-services.asp. (accessed 19 September 2011).

SimulTrans (2011) "Home page of SimulTrans Language Service Provider". http://www.simultrans.com/quality/quality-process. (accessed 5 September 2011)

Tcworld (2011) "Home page of Tcworld Magazine for International Information Management". http://www.tcworld.info. (accessed 10 September 2011)

Temizöz, Özlem (2016a) "Post-editing machine translation output: subject-matter experts versus professional translators", Perspectives Vol. 24, No. 4: 646-665. DOI: 10.1080/0907676X.2015.1119862 Available at http://www.tandfonline.com/doi/full/10.1080/0907676X.2015.1119862

Temizöz, Özlem (2016b) "Counting or not counting rec errors in translation quality evaluation", SKASE Journal of Translation and Interpretation Vol. 9, No. 1: 51-63. Available at http://www.skase.sk/Volumes/JTI10/pdf_doc/04.pdf.