REVISING A TRANSLATION

TIPS AND GUIDELINES REGARDING AN EFFECTIVE TRANSLATION QUALITY CONTROL

"Quality control are measures aimed at checking, measuring, or testing of one or more product characteristics and to relate the results to the requirements to confirm compliance."

CEOpedia

ABSTRACT

Revision is the first and the most important phase of quality control, requiring the knowledge of both the source and the target language, as well as the detailed customer's requirements. The revision process can be computer-aided, yet it is the person performing revision who plays the key role in it, as software used in the process is not able to classify all types of errors, and the ones it does, also need to be reviewed. This publication covers a typical revision scenario, the most frequently used error categories, advantages and disadvantages of computer-aided revision tools, and finally, some tips regarding the use of thereof. It also presents the most frequently mistakes of different types and some hints on how to avoid them.

Proofreading — examination of the revised target language content and applying corrections before printing.

Quality Assurance — a set of activities that allow to ensure compliance of the translated text with the requirements of the customer. Quality assurance is a process performed prior to the proper work with a text, and focuses on preparing a glossary and a set of instructions, as well as on a team selection and answering any queries.

Quality Control — a set of activities that allow to verify whether the text meets a customer's requirements. Quality control is performed after the text has already been processed, and focuses on finding and removing any technical, linguistic, and editorial errors.

Review — monolingual examination of the target language content for its suitability for the agreed purpose.

Revision — bilingual examination of target language content against source language content for its suitability for the agreed purpose.

1   REVISION PROCESS

1.1   Revision vs. Quality Control

According to ISO 17100 standard, quality control should cover three stages – revision, review, and proofreading. Corresponding definitions can be found in the glossary above. Although revision is just

one of three stages of the quality control process, it is, by far, the most important one. Firstly, it is the only stage that is obligatory. Customers rarely provide texts for translation in a form prepared for publishing of any kind, but rather as bilingual files processed in one of CAT tools. In such case, the translator/translation company has no access whatsoever to monolingual files or a prepared publications, therefore they cannot perform any review or proofreading tasks, since these are provided by the customer. Moreover, these stages are not a part of the translation project per se, as they are done by editors, who may as well be unfamiliar with the source language. Considering the above, we have decided to treat revision tasks as an equivalent of the quality control process, yet bearing in mind that this is a simplification of a kind.

1.2   The Role of the Reviewer

Any reviewer working according to ISO 17100 standard is obliged to verify whether the provided translation meets the quality requirements needed for it to be used in the target language. Therefore, the purpose of the text has a significant meaning in the whole process, as an advertising or marketing piece of writing should meet different requirements than a scientific thesis, contract, manual, or an order for a literal translation, i.e. a word-for-word translation.

However, irrespectively of the type of the text, any reviewer is obliged to control whether the translation matches the original by performing a detailed comparison and language revision —to ensure that the text is free from stylistic, grammar, or spelling errors — as well as performing a content-related check. Moreover, the reviewer has to verify that the translator has been working according to instructions and reference material (i.e. style guides, glossaries, etc.) if any had been provided with the source text.

1.3   Step-by-Step Translation Revision

Reviewers should begin their work with familiarizing themselves with instructions and reference material provided with the translation. These should be read as thoroughly as time permits. If the reviewer is not granted enough time to do so, the materials ought to be skim-read at least. (It might be impossible to read carefully a 120-page style guide when the schedule is tight, yet the reviewer should at least know the most important instructions and make sure that all references necessary to fulfill the task were provided.)

Having read the references and instructions, the reviewer is ready to begin the process of revision. The list below covers the optimal steps of their workflow.

- Prepare the text so that both the source, and the translation are visible on the screen. If the translation is prepared in one of the CAT tools, this should be easy, as in such case a bilingual file is at hand. Then, open the file in an application used previously for translation or in any other application supporting said file extension. If the bilingual file was not provided, open the source text and the translation in two neighboring windows, so that the content of both is visible simultaneously.

- Perform an automated quality control using an internal function of the CAT application or an external tool. For details, see *Automated Quality Control*.

- Each segment or a logical part of the source text contained in the bilingual file (e.g. a sentence, paragraph, header) should be read carefully and compared with its translation, with special focus on the following:

  - Is the source properly reflected by the translation?

  - Is the translation complete, are there no parts of the source text left untranslated, and no parts that are not in the source added to the translation?

- Is the language of the translation correct, i.e. is it free from spelling, punctuation, grammar, and stylistic errors?

- Is the terminology used correctly, both in the context of expertise in the field and references, e.g. earlier translations or glossaries; and is it used consistently? (Unambiguity is especially important in technical passages or instruction sets.)

- Is the language of the translation proper for the purpose of the content?

- If there are any tags in the source, are they transferred properly into the target, i.e. are they intact and placed correctly (if the order of the sentence was changed in the translation)?

- It would be wise to focus on particularly important items or such that are especially prone to errors, e.g. proper names, product names, geographical names, addresses, dates, numbers, etc.

- Are all numbers properly formatted (correct decimal and group separators)?

- Are proper quotation marks and dashes incorporated into the translation?

- When the automated quality control tool reports an error, the reviewer is obliged to verify whether this is a valid issue.

- Should that segment or logical part of the text indeed contain an error, the reviewer will take different measures depending on whether the translation is being corrected or just errors to the translator are being reported. If the reviewer has to correct the translation at the moment of reading, which is the most frequent case, the correct version should be introduced in the change-tracking mode. If, on the other hand, only report for the translator, who will later introduce the changes, is to be prepared (which is the best solution, yet scarcely used by translation companies due to lack of time), the reviewer just annotates the error and enters the correct translation, type of the error, and its severity in a proper form (e.g. LISA QA Form).

- Revision performed within a CAT tool requires the reviewer to, additionally, look at the features of the text which would be hidden in a simple word processor, as CAT tools provide also information on the origin of the segment. A translated segment can originate from some previous translation project (so called 100% match) or can be slightly different from it (so called fuzzy match). Of course, it can also be a newly translated segment. This could be really useful for the reviewer as:

  - The decision on what to do with previously translated segments, i.e. 100%-match segments, is entirely up to the customer. In some cases, the customer might wish that such segments should remain unchanged even if they contain errors. Then, changing their content would be against the terms of the order. And likewise, the customer may be aware that the previous translation is of poor quality; then, 100%-match segments need to be thoroughly revised. Moreover, all the in-between situations are also possible, e.g. a detailed instruction to correct only content-related errors, leaving the language of the translation as is.

  - Fuzzy matches are segments containing sentences that vary only a little from the one stored in the translation memory such as in the case of technical documentation where sometimes the only change is a value of a parameter, date, or product name. Fuzzy matches, albeit useful, are extremely error-genic, as the translator can overlook the difference between the new sentence and the one stored in the translation memory, and accept the incorrect segment. That is why fuzzy matches need careful inspection, with particular focus on the numerical values and proper names.

  - New segments (0% matches) contain translation prepared from beginning to end by the person whose work is being revised. If the translator is an inexperienced one, new segments should be revised with extreme care, as it is these segments that show the language skills of the translator.

- The last stage of the reviewer's work is to repeat some steps of the automated quality control (perform spelling, double spaces, number format, and punctuation checks) to make sure that no

typos were introduced to the translation in the review process. This is very important, as nothing discredits the reviewer more than errors left.

- When the work on the text has been finished, the reviewer needs to make certain that the translation is ready to be shipped, i.e. confirms that all the files are in place, in correct format, and free of any comments from the translator, etc.

2   TYPES OF ERRORS

2.1   Error Taxonomy

People often think that preparing a technical translation is a simple and undemanding task because the text lacks the traits of a work of literature. Unfortunately, this is not true. Technical translation is as difficult and complicated as any other type of writing, although it may differ in style and give other opportunities to make errors.

If the technical translations were performed by an expert in the field and a master linguist at the same time, the finished product would be flawless in every way. However, such circumstances are extremely rare, thus it is very important that the translation is checked by an independent reviewer, who is experienced in the field and able to find any potential language errors. It is also worth remembering that the source text may contain errors as well, which is an additional difficulty for the translator, and for the reviewer.

Errors can be divided into the following categories:

- Accuracy-related — which could be further divided into subcategories:

  - Completeness — incomplete translation of source files, including all their parts, i.e. figures and charts titles, tables and numbers in the text or information given in parentheses. A common mistake is to leave untranslated text in the source language.

    Kliknięcie tej opcji umożliwia przejście *to* następnego ekranu.

  - Headers and footers — these should match the titles of the chapters and sections in the translation.

    The title of a chapter translated in a header as *Formatowanie pliku*, whereas the chapter is titled *W jaki sposób formatować plik*.

  - References — improper and inconsistent translation of references to other manuals, chapters, sections, paragraphs, number of pages, etc.

    The reference mentions *page 10*, whereas the topic is discussed *on page 13*.

  - Table of contents and index — subsequent positions in the table of contents and index should match the terms used in the body of the text.

    An entry in the table of contents translated as *Informacje dotyczące pacjenta*, whereas the title in the body of the text is translated as *Wyświetlanie informacji o pacjencie*.

  - Consistency with the source — the translation should always keep consistency with the source, especially with elements such as addresses, phone numbers, model names, manufacturer names, catalog numbers, etc.

    A phrase *Przejdź do następnej strony* translated as *Przejdź do poprzedniej strony*.

    A negative sentence *Nie należy usuwać plików* translated as an affirmative statement *Należy usunąć pliki.*

    A conjunction *or* translated as *i* instead of *lub*.

A catalog number *12345* translated as *1345*.

- Phraseology — improper use of fixed word combinations in a certain context.

    Temperatura dobiegła do maksimum instead of *Temperatura osiągnęła maksimum*.

    Ograniczyć jakość instead of *Obniżyć jakość*.

    Using pleonasms, such as *akwen wodny, wzajemne współdziałanie, poprawić się na lepsze*.

    Using tautologies, e.g. *poprawa i polepszanie zdrowia, pojedynczy i odosobniony przypadek, generalizowanie i uogólnianie problemu.*

    Using oxymora, e.g. *dobra wada, liberalny konserwatyzm*.

- Grammar — not observing grammar rules of the target language.

    Using the accusative case (*kogo? co?*) instead of the possessive case (*kogo? czego?*) with the verb *używać*, e.g. *używać komputer* instead of *używać komputera*.

    Using singular form of nouns that have only plural form, e.g. *dana* instead of *dane*, *igrzysko* instead of *igrzyska*.

    Using plural form of nouns having only singular form, e.g. *wiele oprogramowań* instead of *wiele programów*, *wiele funkcjonalności* instead of *wiele funkcji*.

    Improper use of the demonstrative pronoun *ta*, e.g. *wybrać tą funkcję* instead of *wybrać tę funkcję*.

- Punctuation — not observing the punctuation rules of the target language. Keep in mind that the punctuation rules of the source language may be radically different from the ones of the target language. This also includes inconsistent use of punctuation in activity lists or numbers.

    Inserting space before a question mark, exclamation mark, or colon, e.g. *Usunąć plik ?*

    Improper use of periods in abbreviations, e.g. *nr, dr, itp*.

    Using commas instead of spaces as group separators in large numbers, e.g. *25,500*, *1,000,000*, whereas the proper notation is *25 500, 1 000 000*, or *1 mln*.

    Using a period instead of a comma in decimals, e.g. *0.5, 1.25* whereas the proper notation is *0,5* and *1,25*, respectively.

- Misunderstanding — improper expression of the meaning in the translation, including improper translation of expressions in different contexts.

    The noun *content* translated as *zawartość* instead of *spis treści*.

    The adjective *actual* translated as *aktualny* instead of *rzeczywisty*.

- Spelling — not observing the spelling rules of the target language. Using spelling check function is helpful, but not all spelling errors can be found that way.

    One-word and two-word spelling of various words, e.g. *możnaby* instead of *można by*, *napewno* instead of *na pewno*.

    Using capital letters in the middle of the sentence, e.g. *Wprowadź Wszystkie Informacje* instead of *Wprowadź wszystkie informacje*.

    Omitting spaces between a value and unit, e.g. *500A* instead of *500 A*, *120V* instead of *120 V* or placing a space between a value and the percent sign or the degree of centigrade, e.g. *50 %* instead of *50%*, *36,6 ^{o}C* instead of *36,6^{o}C*.

- Syntax — improper structure of a sentence or improper word order.

    Using the passive voice instead of active, e.g. *to okno dialogowe powinno być wyświetlone* rather than *należy wyświetlić to okno dialogowe*.

Translating the phrase *LCD panel* as *LCD panel* rather than *panel LCD*, *patologiczny stan* rather than *stan patologiczny*.

- Stylistic — changing the style of the translation as compared to the source and reference materials. Manuals and instructions are often written using complicated, technical vocabulary. Such documents should be kept formal, simple, and understandable. More elaborate style is acceptable in marketing-related texts. It is unacceptable to use slang or colloquial expressions.

  Using informal style within formal documents, e.g. *Kochani Przyjaciele* instead of *Szanowni Państwo*.

  Using masculine or feminine gender when addressing the reader, e.g. *Wybrałeś poprawną opcję menu* rather than impersonal form *Wybrano poprawną opcję menu*.

  Using diminutive forms e.g. *myszka* instead of *mysz*, *okienko* rather than *okno*.

  Using foreign words when the target language offers a proper equivalent: *diler* instead of *sprzedawca*, *software* instead of *oprogramowanie*, *flesz* instead of *lampa błyskowa*, *specyfikacje* instead of *dane techniczne*.

- Terminological — improper and inconsistent use of terms included in glossaries, translation memories, and/or other references. Should there be any issues with finding a term in provided references, it would be preferable to use (in the entire translation) the terminology that can be found in public dictionaries.

  Translating *display* differently within the same document, e.g. *ekran*, *wyświetlacz*, and *widok* instead of using a single, dictionary- and/or translation-memory-consistent term.

2.2   Error Weights

Each error can be attributed a weight defining its impact on text clarity. Error weight can be computed on the basis of an algorithm included in a quality control software, which helps the reviewer assess the translation objectively.

Errors can be divided into three following categories:

- Minor error — an error appearing once in the translation and affecting its clarity in an insignificant way. Small typos, negligible punctuation, stylistic, or syntax errors are considered minor.

- Major error — an error significantly affecting clarity of the text, yet in a less distinct manner than a critical error. A text misunderstanding, lack of terminological consistency, stylistic inconsistency, improper context, spelling, punctuation, stylistic and phraseological errors, using phrases not matching the type of text or improper in said context can be considered a major error. A repetitively occurring minor error becomes a major error.

- Critical error — an error affecting the meaning and understanding of the translation in a serious manner. Such errors hinder the user from understanding the functionality of a device or software, or give them false information. Errors appearing in key parts of the text, e.g. on the title page, in the headers or titles, as well as improper translation of a term, a whole sentence or its part, lack of translation, striking spelling, syntax, phraseological, or grammar errors are considered critical. If an error is classified as less significant, i.e. a major one, but it appears frequently in the translation, it becomes a critical error.

Apart from these three categories of errors, there are preferential corrections and suggestions; these are not considered errors, but the reviewer's suggestions related to the specifics of a product or customer's requirements.

3 PRACTICAL GUIDELINES

3.1 Guidelines Concerning the Use of Polish

The major rule of translating technical text is to keep the translation simple and clear. A more elaborate style is required only in marketing-related passages. It is worth remembering to avoid too literal translation and loan translations. The translation should be done correctly, in a manner understandable to the reader. It is a good practice to have the text read either by oneself or somebody else after it has been finished. Several tips regarding Polish are presented below.

To avoid potential errors, it is good to remember to:

- Limit the passive voice in the text:

  Ta wartość nie może być zmieniona.

  Nie można zmienić tej wartości.

- Avoid diminutive forms:

  Należy podwójnie kliknąć myszką.

  Należy dwukrotnie kliknąć prawym przyciskiem myszy.

- Avoid too large adjunct clauses separating the predicate from the subject or the object:

  Ta firma produkuje znane na całym świecie ze swojej niezawodności komputery.

  Ta firma produkuje komputery znane na całym świecie ze swojej niezawodności.

- Avoid redundancies:

  Należy odebrać zamówienia z regionu, zliczyć zamówienia z regionu i załączyć zamówienia z regionu.

  Należy odebrać, zliczyć i załączyć zamówienia z regionu.

- Use impersonal forms when translating software dialogs:

  Nie mogę zachować pliku.

  Nie można zapisać pliku.

- Avoid using specific grammatical gender (masculine or feminine). Rather, keep the language gender-neutral:

  Powinieneś wprowadzić tę wartość.

  Należy wprowadzić tę wartość.

- Avoid addressing the reader with a possessive pronoun, such as *Twój*, *Twoja*, or *Twoje*:

  Uruchom ponownie Twój komputer.

  Uruchom ponownie komputer.

- Do not use abbreviations at the beginning of the sentence:

  Np. można użyć tej opcji.

  Na przykład można użyć tej opcji.

- The verb *użyć/używać* conjuncts with the second (*kogo? czego?*), not the fourth grammatical case (*kogo? co?*). The verb *zapobiegać* answers the question *komu? czemu?*

  Użyj opcję, aby wprowadzić wartość.

  Użyj opcji, aby wprowadzić wartość.

Rozwiązywanie i zapobieganie problemów.

Rozwiązywanie problemów i zapobieganie im.

- The proper form of the word *ta* in the accusative case is *tę*, not *tą*, which is the proper form of the instrumental case:

    W polu należy wprowadzić tą wartość.

    W polu należy wprowadzić tę wartość.

- Use comma, not a period as a decimal separator:

    Kalibratory mogą mieć średnicę: 3.1, 4.0 lub 4.9.

    Kalibratory mogą mieć średnicę: 3,1, 4,0 lub 4,9.

- Use Polish convention of grouping in large numbers:

    Łączna kwota wynosi 1,500,200 zł.

    Łączna kwota wynosi 1 500 200 zł.

- Distinguish countable and uncountable nouns:

    Ilość godzin.

    Liczba godzin.

- Inflect words like *agent*, *kreator*, *klient*, and *menedżer* animatedly, i.e. *agenci*, *kreatorzy*, *klienci*, and *menedżerowie* when talking about people, and inanimately, i.e. *agenty*, *kreatory*, *klienty*, and *menedżery* when talking about inanimate objects such as software.

- If an acronym is present in the sentence, always use the grammatical gender of the main noun of the abbreviation. Do not use the neutral gender to all acronyms, e.g. *IP* is *Ten IP (adres)*, and not *To IP*.

- Remember to differentiate the elements of the UI from the parts of a keyboard or computer mouse. If *buttons* appear on the screen or are parts of a pointing device (e.g. *OK*, *Usuń*, or *Dodaj* or *prawy przycisk myszy*), they are *przyciski* . If they are parts of the keyboard, they are *klawisze* (e.g. *Delete*, *Insert*, or *Enter*). One can only *nacisnąć* a button or a key, never *wcisnąć*.

- Never leave the *etc.* abbreviation untranslated. Its Polish counterpart is *itd*.

- Copyright-related terms should be translated as follows:

    Trademark should be translated as *znak towarowy* rather than *znak handlowy*.

    Servicemark should be translated as *znak usługowy*.

    Registered should be translated as *zastrzeżony* rather than *zarejestrowany*.

    Copyright should be left as is.

    All rights reserved should be translated as *wszelkie prawa zastrzeżone*, not *wszystkie prawa zastrzeżone*.

- The phrase *Note* should be translated as *Uwaga*, *Caution* as *Przestroga*, *Warning* as *Ostrzeżenie*.

- Colloquial English phrases such as *you should*, *you had better*, *you may want to* ought to be translated as *można*, *należy*, or *istnieje możliwość*.

- The reader should be addressed impersonally, e.g. *należy*, *można*, *wykonano*, etc. Polite addressing, e.g. *Państwo*, is accepted in formal letters or official marketing papers.

3.2 Guidelines Concerning Quality Control

There are several routine, yet mandatory operations that help eliminate errors and lead to increased quality of the translation. These actions should be taken during the translation and after it has been finished.

While translating, it is worth to remember to:

- Use the references delivered with the source text.

- Follow the instructions exactly.

- Use consistent terminology in the text.

- Do not change the sub- or superscripts, nor the ™, ®, and ° symbols.

- Do not remove, add, or change the tags; at the same time, remember to put the context tags such as ** in the part of the translation that matches the meaning of the emphasized source.

- Keep to the guidelines, especially those delivered by the customer.

When the translation has been finished:

- Check if it is complete.

- Check the spelling.

- Check if the subscripts and superscripts, as well as the ™,®, and ° symbols are in the right places and formatted correctly.

- Check if all the parts of the text are translated and positioned as they should be: the headers, footers, captions for figures, tables, and charts, as well as the list elements.

- Check if the translation is stylistically consistent.

- Read the whole translation or ask someone to do it.

4 AUTOMATED QUALITY CONTROL

Automated quality control tools are invaluable help for the reviewer. They import a bilingual file created by a CAT application, and compare the source against the translation according to specific criteria. Such automated quality control allows to detect many errors arising due to human carelessness or technical issues. These are not content-related, grammar, or stylistic errors, yet they are the most frequently found in any translation.

Most of the automated quality control tools detects the following types of errors:

- Improper spacing, e.g. double spaces or additional spaces at the beginning or end of a segment.

- End punctuation differences, e.g. period in the source and lack of thereof in the translation.

- Additional parentheses or lack of parentheses in the translation.

- Spaces before specific punctuation marks.

- Target the same as source, i.e. lack of translation whatsoever.

- Empty target segments.

- Differences in capitalization at the beginning of the source and target segments.

- Inconsistencies in the translation, i.e. the same segment translated in several ways or different source segments translated identically.

- Discrepancies in dates, numbers, and units of measurement.

- Discrepancies in segments length. (These tools usually allow to define a maximum allowed percent difference between the source segment and the target segment.)

- Conformity with the glossary. (Most of the tools are incapable of proper inflection handling, therefore the use of that particular function is limited.)

- Conformity in the use of tags and other formatting characters, e.g. end-of-line character, tabulators.

- Spelling in the target language.

The results of automated quality assessment can be later reviewed interactively or exported to one of multiple file formats.

It needs to be stressed that the automated quality control software, however useful it might be, is only an auxiliary tool. Such applications cannot think like men, therefore they are not to be used as a replacement for the reviewer. Firstly, none such software can detect content-related, grammar, or stylistic errors. Secondly, it is not uncommon that the issues reported by these tools are not errors at all. Such "false errors" appear e.g. when a number is written in words in the target language, capitalized words from the source are not capitalized in the target (like in the case of German nouns), or a segment contains an untranslated proper name. Hence, an automated quality control report has to be always checked by a person. Moreover, it is often unnecessary to use all the available options in the tool.

Here are examples of situations when some of the automated quality control options are not worth using:

- Performing a terminology check in inflected languages such as Polish — a word appearing in the segment in a grammar form different from the basic one will be always reported as an error.

- Looking for untranslated segments in texts containing untranslated names or symbols — each cell of a table is a separate segment in a bilingual file, therefore the application will report an error for each cell containing a name or symbol.

- Checking capitalization compliance in German — all nouns in German begin with a capital letter, so many of the errors will be false.

5   SUMMARY

The reviewer is the most important person of all taking part in the translation project who actually influences directly the quality of the finished product. Not only has the effective reviewer to know the source and target languages, the field of the translation, and the requirements of the customer, but also must be proficient in using the automated quality assessment tools, as well as classifying and weighting the error. The decisions made by the reviewer are usually final. Although there are several computer solutions designed for automated quality control purposes — both standalone and CAT-implemented — their application is limited as there are much more possibilities of making a mistake than any of the programmers might have ever think of. That is why the companies aiming at high quality of translation must hire qualified reviewers and keep developing their skills.