

TRANSLATION QUALITY: MODELS AND TOOLS

Marek Pawelec

Everybody wants their translation to be top quality, and subconsciously we all “know it when we see it”, but what does it actually mean? How do you define the quality of a translation? In this presentation I will show some theories on what constitutes “good” translation and how to assess the quality (without going too far into academic debates). But theories aside, there are established, widely used models of the translation quality assessment used in the technical translations industry, such as LISA, J2450 or TAUS error typology. I will show the basis of these models and how they are used in the everyday functioning of thousands of translation agencies across the world. I will also discuss the tools available for translators and translation agencies for quality assurance, both stand-alone and incorporated into translation environments or CAT tools, comparing included checks, error-resolving procedures and customizability. Tools for quality assessment will be discussed too.

Introduction

The translation is a complex process in which a translator subjectively interprets the intentions of a source text writer, trying to recover author’s communicative aims. There are numerous concepts of the translation quality, and many of them are based upon subjective impressions. Although they might offer some rules and guidelines regarding the assessment, in general they still can’t provide a quantitative way of evaluating a translation. One might say that when it comes to quality of translation, we usually still agree on the famous Justice Potter Stewart

quote from 1964: “I know it, when I see it”. The quality of a translation is also strongly connected with the type and quality of the source text. When dealing with the translation of a poem, a play or a novel, the expectations for the translation are usually quite different than in the case of a service manual or a product catalogue. Since the subject of translation quality in case of literature or poetry is still largely a matter of scholarly debate, it won't be discussed here. Instead the presentation will focus on the subject with strong practical and business implications – the translation of the technical texts.

Technical text

What is a technical text? In general, text may have many functions: to entertain, to enlighten, to propagate ideas. The technical text is a kind of text the main function of which is to instruct, but also to describe and to argue. Instructive texts tell the reader how something should be done, descriptive texts – how things are. As you probably know from experience, these kinds of texts are usually found in all types of instructions, user manuals, reference guides and similar publications. Since they serve certain functions, they usually are (or should be) written with respect to certain rules and – ideally – with controlled language, which is a subset of natural language, obtained by restricting the grammar and vocabulary in order to reduce or eliminate ambiguity and complexity. Those qualities make it easier to develop certain methodologies for assessing the quality of the translation of such texts, since it is relatively easy to establish the equivalence of the target text (TT) with regard to the source text (ST). TT should ensure similar efficiency with regard to the syntactic, semantic and pragmatic function of ST within the cultural frame and expressive potentials of both source and target language. Since, if we accept the pragmatic textual approach, that translation operates not with sentences but with utterances, the equivalence should be reached at the pragmatic level. Even when we take into consideration all these factors, it is still hard to define a translation quality (TQ) in a way which would make it possible to measure it quantitatively.

Translation quality assessment

Why measure the translation quality? Especially, why measure it quantitatively? On the current, very competitive market of translation companies, the quality can – and is – a very important factor for both clients and translation providers: offering a higher quality of translation can lead to more jobs and higher income. However, it is very hard to

improve something one can't measure and express in numbers, hence the need for method of translation quality assessment. Metric provide a way to objectively quantify a process and reduce costs of poor quality. So, is it possible to measure TQ? Yes, although indirectly, thanks to the quite simple trick. While it is hard to agree on what a good quality translation is, it is much easier to agree on what is an error, and by common definition an important factor of quality is the absence of errors. Based on this simple idea the practical definition of good quality translation was reached: a good translation is a translation with very few or no errors at all. When measuring TQ we really measure the incidence of various types of errors and defects in the translated material:

- errors of form,
- errors of meaning,
- errors of compliance.

Studies in this field led to the development of the LISA quality assurance (QA) model. LISA stands for Localization Industry Standards Associations – the body which unfortunately has shut down in 2011 due to insolvency, which was the result of a lack of the clear view of its goals. However, during its functioning LISA mandated many standards important for the translation and localization industry and one of them is the LISA QA model, which is a method of assessing the translation quality widely used in the industry.

All this resulted in the development of the quality control system based on Translation Quality Index (TQI) methodology. The TQI methodology is a quantitative-based method of translation quality assessment. It measures the number and type of errors found in a text and calculates a score, or TQI, which is indicative of the quality of a given translation. TQI indicates the quality of a given translation sample, and is obtained by the assessing a translation (usually a random sample with word count greater than a minimal value, usually about 1000 words) by an evaluator with rigorous application of a quality assurance methodology. TQI attributes a numerical value to a translated text, with 100 being an “error-free” translation. The value of a text is based on the number of error points in a given text or sample.

What are the error points and how do they differ from errors? Not all errors are equal: for example, there's a difference between a typo on the front cover of a manual and the same typo in a footnote. This leads to assigning different weights to errors depending on their consequences. In the previous example minor typo can be assigned the weight of “1” and major typos will be scored with some greater weight. Those weights are called “error points”. There's a separation between the error type and its

severity – each error can be marked as critical, major or minor, depending on its consequences.

To achieve certain objectivity, the translation assessment procedure itself has to meet certain criteria. The translation quality measurements should be:

- Repeatable – two assessments of the same text should yield similar results.
- Reproducible – assessment of the same text by two evaluators should give a similar result.
- Objective – void of subjective bias (this requires rigorous selection criteria for evaluators).

Of course a lot depends on the evaluator – good evaluator must be able to be as objective as possible, and be able to distinguish between factual, tangible errors and stylistic preferences. Differences in stylistic preferences are not errors and are ignored in the computation of the quality score, so it is necessary to establish clear rules that define what an error is and what it is not.

When assessing a sample text the evaluator has to answer three questions:

- Is the translation accurate?
- Is it grammatically correct?
- Is the translation compliant with the glossary, style guide, guidelines and client instructions?

If the answer to these questions is “yes”, it means that there is no error.

Of course the TQI methodology cannot measure everything. Since the TQI methodology is designed to measure tangible, factual errors only, it is ineffective when a high degree of creativity is expected on the translator part, like in case of literary, marketing and advertising texts and is best suited to the assessment of technical texts.

LISA standard defines several categories of QA errors:

- Mistranslation (incorrect understanding of the source)
- Accuracy (omissions, additions, cross-references)
- Terminology (glossary)
- Language
 - Grammar
 - Semantics
 - Spelling
 - Punctuation
- Style (adherence to the style guide)
- Country (country standards, e.g. numbers separators, local suitability)

- Consistency

Each error falling into one of these categories can be classified as minor, major or critical. In the standard LISA model minor error have a point value of 1, major 5 and critical is a maximum error points plus 1. When reviewing a translated text or translation sample of volume sufficient to avoid statistical bias, a reviewer enters each encountered error either into the specially prepared Excel sheet (Figure 1), or into special software for TQI calculation, like *aliquantum* or *LISA QA Model* (Figure 2). The sheet or software calculates TQI value based on the number of errors in each category. Using this values and the number of words of evaluated text the software calculates TQI value, which is the basis for the overall classification of the translation. Usually to “pass” the evaluation the text does have to reach a score of at least 94 (out of 100), but the scores in the range 94–98 points are considered inferior and – depending on the contract – incur some financial penalty for the translator. If the text is scored below 94 TQI, the translation is rejected and the translator may not receive the payment.

Of course in practice different companies dealing with large quantity of translations, like some large software and hardware manufacturers and some global-scale translation companies, developed their own, individualized approach. In some of them it was decided to assign fixed weights to certain error categories – for example classifying all terminology errors as major and all punctuation errors as minor. Some versions of the LISA model classify all Mistranslation and Accuracy errors into one category, since the result is basically the same in both cases, and usually this type of errors is regarded as critical. Some implementations of this methodology are adding also a “formatting” category, which in the LISA QA sheet is scored separately. Formatting error can often be classified as major or critical, since the erroneous placements of formatting tags can lead to high cost of DTP troubleshooting. It’s worth noting, that in some cases category “Critical” includes any type of error which may result in injuries, product recalls or other considerable costs, so in an extreme case even a misplaced comma can be classified as a critical error. And presence of a critical error may – or, in case of some companies must – disqualify the translation, which means that it have to be returned to translator/translation company for through review, and financial penalties apply.

Quality assurance features of computer translation tools

The translation market, in general, consists of three groups: end clients, translation companies acting as intermediaries, and actual translators. Thanks to LISA QA model, translation buyers have tools for assessing the quality of technical translations. If a translator wants to deliver a product that will be accepted and classified as of a good quality, first of all he or she does have to know the rules by which his/her translation will be judged by. This is the key message of this publication – to be able to succeed in the world of technical translation, one does need not only the knowledge of a language (especially the specialized, technical terminology of the field he or she is going to work in), but also the rules which are being applied in the industry.

Of course, it is not enough to know the rules – one has to follow them. Most of the technical translation is currently being done with the help of the specialized software – Computer Aided Translation (CAT) tools, which help to streamline the translation of repetitive text by utilizing a database – called translation memory (TM) – containing translation of all sentences (segments) translated previously in a given project.

In order to help their customers, companies offering CAT tools introduced into them many functions which help to reach error-free translations. Those are known as quality assurance (QA) functions, designed to help to eliminate many technical errors, occurrence of which may lead to disqualification of the translation. Of course the QA checks can only help with certain error categories – like terminology, consistency, formatting, spelling or country standards – and are powerless against others, like accuracy or style.

memoQ is one of the “top tier” CAT tools used by translators and translation companies, and its QA functions are on par with similar software products. memoQ offers comprehensive list of checks in the following categories:

- Translation consistency: whether identical segments are translated the same way. The function can control source to target consistency (identical translation of the same source texts) and target to source consistency (identical translation of different source texts).
- Terminology use: if a glossary is available, the software verifies if the correct terms were used. Optionally the “forbidden” terms may be defined.

- Segment length: verification of the translation text length as compared to source text length – length limitation may be critical in a software localization.
- Completeness of translation: not translated segments or those with translation identical as source.
- Formatting: if the target segments contains source formatting elements (bold, italic, underline).
- Numbers verification: whether translated segments contain the same numerical values with proper, localized formatting.
- Punctuation: ability to detect non-standard punctuation marks, missing brackets, quotation marks and incorrect use of spaces after punctuation marks.
- Spaces and capitals: detection of double spaces, additional spaces at the end of segments and verification whether target segment uses the same capitalization as the source.
- Inline tags: whether all the tags (usually formatting elements) from source segment were inserted into target text.

Trados Studio
 Wordfast Pro
 ...

Specialized Quality Assurance software

Since the QA functions of CAT tools are relatively new feature, and not all programs offer the same checks, there are independent software tools designed especially with this functionality in mind. Most of them are commercial, like ErrorSpy (D.O.G. GmbH), QA Distiller (Yamagata) or Verifika, but there is also a free desktop tool, ApSIC Xbench, and even some free online tools for QA. ApSIC Xbench was developed by Spanish translation company ApSIC Localization Solutions and serves two main purposes: one is a front-end for glossaries in many different formats, the other, is a QA tool. To perform its functions the software needs a bilingual files, i.e. files with both source and target text for each segment in one of the many formats supported by the program – amongst more popular are Trados bilingual DOC files (see Figure 4) and XLIFF files (XML Localisation Interchange File Format).

There are several QA checks grouped into four categories:

- Spellcheck (the spellchecking using the open-source engine Hunspell).
- Basic (untranslated segment, inconsistency in source, inconsistency in target, target same as source).
- Content (tag mismatch, numeric mismatch, double blank).
- Linguistic (key term mismatch, project checklist, personal checklist).

[Other tools ...]

Main differences between CAT tool specific QA and a specialized program like Xbench is the ability of the latter to process and verify files from many different CAT tools, usually offering additional checks, not available in some particular CAT tools. For example, the unique feature of Xbench is the ability to execute user-defined checks based on regular expressions, which enables enforcement of some grammar or punctuation rules or identification of stylistic errors.

[LQA tools]

Summary

To summarize: while scholars may still argue for years to come on the concepts of good and bad quality of translation, the translation and localization industry developed tools for quantitative assessment of the quality of translation of technical text involving the relatively simple method of tracking and evaluating errors – all based on the assumption that a good translation is the one without errors. One of the goals of linguistic education is to prepare students for different careers they might take up, which might include a technical translator position. To achieve these goals students must learn not only languages, but also at least basics of the trade they might wish to develop. Since knowledge of QA methodology and criteria used in the translation assessment procedures is an important factor of individual translator success, I believe that they should be presented as a part of the future translators curricula, together with tools and methods for verifying and improving the technical translation quality.

Marek Pawelec graduated in Molecular Biology at the Jagiellonian University in 1992. Worked as the researcher at the Faculty of Medicine and Faculty of Chemistry, Jagiellonian University. Since 2001 working full time as an English to Polish freelance translator, translating literature. Thanks to education and hands-on experience was able to extend his services to translation of medical, biochemical and chemical texts. Translated 30 novels and several millions pages of medical texts, mainly medical equipment manuals. Experienced computer aided translation (CAT) software user and trainer. Certified memoQ trainer. Presented lectures on several translation conferences. Teaches on practical aspects of freelance translation, CAT tools and translation related software.