

QUALITY ASSESSMENT AND ECONOMIC SUSTAINABILITY OF TRANSLATION

Luigi Muzii
Gruppo L10N, Roma

Abstract

The concept of quality is mature and widespread. However, its associated attributes can only be measured against a set of specifications since quality itself is a relative concept. Today, the concept of quality broadly corresponds to product suitability – meaning that the product meets the user’s requirements. But then, how does one know when a translation is good? No answer can be given to this very simple question without recall to translation criticism and the theory of translation. However, the relationship between a source text and the translated text is unfit to solve the problem, as readers often perceive the end-product of translation as the only material available for scrutiny; they have no interest in the translator’s decision-making process (the hermeneutic process). Therefore, translation adequacy should be taken into account in assessment especially when customers impose their own subjective preferences (requirements).

Whatever is worth doing at all is worth doing well.
Philip Dormer Stanhope

Although historically mature and widespread as a body of principles, the concept of quality regarding the production and the delivery of services has dramatically transformed over the last quarter of century to become a relative concept that broadly corresponds to product suitability. But quality is also about customer satisfaction, work efficiency, team working, control and communication. The pragmatic approach in translation studies has helped by considering translation as the product of a process depending on specific expectations, needs of the target audience and its function in a given context or situation. Functionalism makes the traditional notion of linguistic equivalence obsolete and makes functional equivalence more relevant.

1. Requirements defined

General criteria are necessary to standardize the production process and appraise quality in the sense of the product's ability to meet requirements. When dealing with quality, two basic principles must be acknowledged:

1. quality is relative: people can perceive different quality levels in the same product;
2. quality levels are subject to constraints in requirements.

The fundamental assumption in quality standards (namely the ISO 9000s) is that business processes can be improved so that the product pass as it is. For business processes to produce the expected outcomes, the following four elements are necessary:

- basic skills for task completion;
- appropriate and correct information about the job;
- accurate and suitable tools and materials to fulfill each task;
- a well-suited environment.

When these elements are all available, their effectiveness can be measured and possibly improved; controls can be reduced to a minimum, and savings will be at least equal to the planting cost of the whole system.

1.1. Specification of requirements

A specification of requirements is a document providing an adequate and unambiguous description of the task for a project, together with a description of the desired results, the essential conditions to which the service must conform and the characteristics or features of each deliverable.

In a mass-production environment, most products meet most of demands, but leave many real desires unfulfilled. There are plenty of choices, but almost none precisely matches expectations, so buyers are used to settling for less, but do not stop wanting something more. What usually happens is that goods and services are offered to perform tasks or meet needs. But if customers' expectations were actually delved into, it would be discovered that expectations concern transformation. Customers expect the things they buy to make them different. What is pretty obvious with personal items is just as true for business decisions. This deep desire only tends to emerge after needs are met. Understanding and satisfying this desire creates loyalty, and customer loyalty is perhaps the most important element in any product's long-term success.

Quality is essentially conformity to requirements that come primarily from customer needs. In other words, what the customer says is quality is quality, even though meeting the requirement does not necessarily mean producing quality: one could meet all the customer requirements and still produce junk.

Most ‘quality problems’ in translation have little to do with mistakes, and more to do with a mismatch of assumptions and goals between the people requesting a translation and the people supplying it. Also, it is not always a straightforward task to gather requirements from the user. On the other hand, if you can’t collect the requirements you don’t know your customer; and if you don’t know your customers you can hardly expect to please them. Simply stated this means that if a translation cannot be used to accomplish the task it was required for, it has no real use and belongs at the bottom of the cat box. This is why academic disputes about quality are useless in practice: no customer will be willing to spend time to get involved.

The key to quality translation is really the ability to successfully negotiate between competing demands to find the translation that fits a particular situation, and represents the best trade-off between requirements that cannot all be simultaneously met. The name of the European quality standard for translation services EN 15038:2006 reads “Translation services – Service requirements”, and its purpose is to establish and define the requirements for the provision of quality translation services. Admittedly, a key issue is quality assurance and the ability to trace its progress. Nevertheless, despite its efforts, the Italian delegation did not succeed in obtaining a commitment about Service Level Agreements (SLAs) and metrics being included in the final draft. A service-level agreement is a contract between a service provider and a user of that service (the customer) that specifies the level of service that is expected during the term of their agreement. It also defines the terms of the provider’s responsibility to the customer and either the type and extent of remuneration if those responsibilities are met or the extent of penalty if they are not met.

The lack of specification of any translation quality metrics is a serious *vulnus* when assessing the process of a translation service provider ready for certification to the new standard. Anyway, in 5.2.3 Linguistic aspects, the CEN standard requires

that information about any specific linguistic requirements in relation to the translation project is registered. Such information can include requirements of compliance with a client style guide, adaptation of the translation to the agreed target group, purpose and/or final use, use of existing terminology, and updating of glossaries.

Different types of documentation need different quality requirements. Owner’s handbooks need to be easy and enjoyable to read as well as being technically correct – strange as it may sound, there are people who actually read them. Workshop/repair manuals need to be technically correct, but style is not so important as long as it’s understandable. Most service technicians will only look up the procedure they are interested in and they only need to understand

the steps they need to carry out. A mistranslation that causes the reader to misunderstand or carry out an operation incorrectly is a serious mistake, a failure. A stylistic error in a workshop manual is a minor error, but is a more serious error in an owner's handbook. This is what the expression "fit for purpose" essentially means and explains why different metrics should be used for different types of texts.

1.2. Metrics

Metrics are a set of rules that allow users to measure how much a product (the translation) meets requirements, and are generally used to measure performance. The primary goal of measuring, of course, is to create a standard against which something can be judged. What is often forgotten is that metrics can be used not only to measure performance, but also to identify specific problems that are affecting performance. Long before Heisenberg developed his uncertainty principle, it was well known that the act of measuring influenced the system being measured. Also, measuring serves little purpose if it provides no means for improvement. Therefore, when developing a metric the aspects of quality everyone will work to improve must be defined.

Effective metrics must be objective (measurable), unbiased, and able to provide enough resolution (detail) to assess the factors that need improvement. This means that any two people who set out to calculate the value of a metric must be able to produce comparable results. Typical metrics are SAE J2450 (recently elevated to standard) whose goal is to provide "a tangible method for measuring the quality of translation deliverables as precisely as for any manufactured product" (SAE 2005). SAE J2450 provides for minor and severe occurrences of wrong terms (glossary violation or conflict with *de facto* standard translations), syntactic errors, omissions, word structure or agreement errors, misspelling, punctuation errors, and any linguistic errors related to the target language that are not clearly attributable to the other categories.

Subjective metrics are hard to measure because their value depends as much on opinion as on demonstrable facts. Translation quality is a typical case of subjective assessment. All translations are prone to subjective influences due to the subjective conditions of the hermeneutic process and the translator's personality; and reviewers and editors are subject to the same influences. Hence quality is always a very personal issue, a relative matter. Perception is everything. This also explains why translation quality is such an endlessly-debated subject causing fierce and divisive disputes. There are those who claim that the only key to 'quality' translation is some form of certification or accreditation scheme for translators based on academic qualifications – or equivalent – and generally combined with membership of a 'professional'

organization, and there are those who argue that consistent and acceptable translation output quality can be achieved most effectively through quality-oriented process design and standardization, possibly supported through common standards.

The first argument is increasingly suspected to be based on the desire to limit access to the profession to an elect group of ‘professionals’ meeting criteria that they themselves have devised. The second argument is equally suspected as being flawed: it is impossible to create any metrics of quality assessment due to the substantial amount of craftsmanship, creativity, and subjectivity involved in any translation. Not surprisingly, the ivory-tower conception of translation is midway between science (translation science) and art. It produces thousands of ‘graduate translators’ emerging onto the market every year, confident in their in-built superiority and ability to provide ‘perfect’ translations, yet quite unprepared (for the most part) for the harshness of an increasingly savage competition.

If translation is a science, translation assessment should be as well. Words are like stones, but translation theorists seem to deliberately forget this long-life simple principle. How much can Galileo’s principles on experience be applied to translation?

From the user’s perspective, the assessment of a translated text should be made regardless of its nature, and the translated text should be considered simply as a primary text. Should any other approach be considered as valid just because translators are so fond of themselves and of their job? In some respect this question seems to have more than something to do with the frustration of doing a job that is poorly appreciated both in social and economic terms: translation is undoubtedly one of the least remunerated jobs that can be offered to any individual with specific cultural requisites.

On the other hand, translators don’t like being told about their errors. This idiosyncrasy can be put down to human nature, an instinctive hostility to criticism, and to the importance that translators give to their job for the mental effort that they lavish or think they lavish on it. Clearly, this attitude is prejudicial for any objective approach to quality, and is often deemed as impossible.

2. Translation quality assessment

The definition of “quality” as stated in ISO 8402:1994 (ISO 1994) reads: “the totality of features and characteristics of a product or service that bear on its ability to satisfy stated or implied needs” (3.1). Quality is also defined as an integration of the features and characteristics that determine the extent to which output satisfies the customer’s needs. Both definitions implicitly depict the

customer as the best judge of the quality of a translation, which is true as long as he has the capacity to dictate the stringent requirements for the service. These requirements – stated or implied – play a central role, and will eventually be expressed in terms of attributes. In ISO 8402:1994 “a defect” is defined as the non-fulfilment of intended usage requirements (3.21).

The refusal to introduce SLAs and metrics in the EN-15038 European standard lies in the belief that, generally speaking, the customers of a translation service do not have the necessary skills and competences to drive the provision of service through requirements and that, in practice, they rely on the service provider to deliver a certain degree of intrinsic quality. As there are virtually no tools available to validate compliance to standards – however unstated – the refusal of metrics is a direct consequence.

Since there is no ‘perfect’ translation, the intended purpose of a translation and its suitability remain the only judgment criteria that, for the sake of objectivity, should be accompanied by assessment metrics. The combination of process and output quality assessment of a translation can only show whether it is acceptable or defective. So, translation quality assessment (TQA) criteria have to be agreed upon with the customer, be a subject of requirements and be formalized in a separate document.

So far, TQA has been performed on the basis of strict correspondence between source and target texts and on intensive error detection and analysis. While this is undoubtedly the best approach from a theoretical – and maybe pedagogical – point of view, it is totally uneconomic as it requires a considerable investment in human resources and time, and reduces translation to a matter of trust – which unfortunately is also current practice – since no technical translator trained by current university teaching methods and programs is properly prepared to meet different quality criteria.

Assuming, then, that it is impossible to set objective ‘aesthetic’ parameters for quality translations, it is quicker and easier to formulate a generally negative judgment based on whether proper equivalence of signs exists between the source and the target texts. Conversely, when a customer or a reviewer rejects or dislikes a translation, three steps should be taken – something that rarely happens before taking on the job:

1. arrive at a full understanding of the linguistic quality requirements of the customer;
2. agree with the customer on a process to correct any deviations from requirements;
3. implement a process to prevent the same issues from occurring in the future.

Basically, linguistic quality consists of five components:

- correctness
- completeness
- meaning
- terminology
- style

Meaning can be traced by comparison: translation should allow its user to perform the same task as the original piece of text, which is almost impossible when the two texts ‘read’ differently. This more or less explains why style is much too often the prime cause of dissatisfaction with a translation. On the other hand, every translator makes his/her own choices that become apparent in any deviations from the source text; a poor translator is not the one with a questionable style, but the one with no style at all.

Terminology is the second component, as translators unfortunately do tend to switch terms even if they have been instructed not to. In fact, many translators follow a code of creativity that might read as follows:

1. I can write it better;
2. if I can find a better term than the existing one, I will use it.

In reality, who will check 100,000 words of translation to find any terminology changes after the translations have been delivered? However, if terminology issues can be approached in a systematic way, style can then be left as a matter of personal preference. The same goes for correctness and meaning with respect to completeness. Any translation can be roughly checked with the source text for completeness. Yet, correctness requires a specific knowledge of grammar, spelling, and punctuation, and it is understood as conforming to an approved or conventional standard, freedom from fault or error. These points, however, are quite often taken for granted when the job is done by a professional translator.

A detailed statement of the work to be done and an accurate style guide, although time consuming, can be helpful in most situations – possibly together with examples of do’s and don’ts. Especially for large projects, translation should, and now could, be considered as a production process, by the same standards of common business. In this perspective, defects should be able to be reproduced under the same conditions, corrected and then removed. This approach would eventually lead to set defect tracking and assessment procedures, and to pass/fail criteria for sample testing.

2.1. Measurability

Sir William Thompson, first Baron Kelvin, in his lecture to the Institution of Civil Engineers of May 3, 1883 stated:

When you can measure what you are speaking about, and express it in numbers, you know something about it; but when you cannot express it in numbers, your knowledge is of a meager and unsatisfactory kind; it may be the beginning of knowledge, but you have scarcely in your thoughts advanced to the state of science.

In the language industry ‘quality’ is one of the most debated subjects: it is part of daily conversation. The cursed triangle of time, pricing, and turnaround time seems to take up the whole agenda. The most commonly-asked question about quality is: how can quality be measured? To measure something, you must know what it is, and then you must develop the metrics to measure it. Metrics definition is the hardest part for people who have always thought of quality in their deliverables as a questionable subject. The best way to assess quality remains that of measuring the number and magnitude of defects; and when defects cannot be physically removed, their features and scope must be specified. In this respect, translation quality can obviously be assessed by comparison with the source text, but if a flawed translation is quite easy to detect, at least in terms of its ‘suitability of purpose’, the quality of a fair or good translation will often be called into question by external factors such as personal taste.

The first step, then, is to establish a model or definition of quality, and translate it into a set of metrics that measure each of the elements of quality in it. Measuring things just because they can be measured, though, is not useful. If something is not relevant to the quality model established, it is not a good use of time to develop metrics to measure it.

Striving for a single, all-encompassing metric is not only troublesome, but it can also be useless, as a simple metric would not reveal all the problems. Creating multiple metrics that assess the various aspects of what is to be measured can help re-compose the overall framework: knowing which parts of a process work well and which ones don’t allows for taking measures to correct the problems. A comprehensive set of metrics must measure quality from several perspectives and at several points during the production process, regardless of the quality model. At a minimum, metrics should tell something about:

- quality of the finished product;
- lack of quality of the finished product;
- quality of the process – how reliable it is to produce quality products;
- likelihood of achieving quality in this deliverable (predictors of quality).

The quality of the finished product corresponds to general customer satisfaction ratings, while the lack of quality can be measured by defects such as technical errors. The quality of the process comes from repeatability, and typical predictors of quality are in-process indicators such as editing.

Levels of translation quality can be described at least in the following terms:

- discard
- raw
- standard
- finished
- adaptation

Raw translation means a translation that conveys the central meaning of the original text. There may be grammatical errors and misspellings, but the text will be understandable. Typically, this level could apply to translations of large amounts of scientific abstracts. Standard translation corresponds roughly to the translation of antiquity. The original text is translated fully and the translated text is grammatically correct and reasonably fluent. The text may be awkward at times, but the contents of the original text should be understood completely from the translation. Typically, this could be the translation of a technical manual.

A finished translation implies that the translated text is both fluent and idiomatic, and could be assimilated completely to the cultural context of the target language. One should not be able to recognize the translated text as a translation. Typically, this could be an advertisement brochure or a piece of literature.

Adaptation is not actually the direct translation of a text but the production of a new text based on foreign language original(s). The resulting text need not correspond sentence by sentence to the original(s), but may instead even have omissions or re-orderings according to what the translator deems appropriate. The language of the resulting text is expected to be fluent.

Most quality components can be clearly described and precisely verified. Again, what makes language so elusive is its subjective nature. Thus individual habits and preferences far outweigh academic considerations for all practical purposes. People can become extremely passionate about their preferences, down to endless rounds of revision and pointless debate. So translation providers cannot really guarantee linguistic quality without any input from the people who will ultimately judge this quality. In other words, to have firm control over linguistic quality, the relationship between producers and users – the rules of engagement – must be defined, implemented and adhered to.

2.2. Rules of engagement

Because quality is so subjective, and its definition is such a relative thing, developing quality specifications for each new project is a good method for clearly setting quality parameters. However, determining the accuracy of a text is a highly intellectual and creative skill, and the customer rarely has the

knowledge of the quality necessary to lay down specifications by allocating the necessary resources to produce what s/he will eventually be happy with. Therefore, going beyond the customer's requirements to produce what is deemed of high quality always implies allocating one's own resources.

Translation quality should be tracked from different perspectives: number of reviews and time spent on each of them, number of errors found, productivity, and suitability. Being able to track translation defects is not only an important condition for delivering high-quality services to customers, it also provides an efficient way to evaluate vendor performance. In fact, the reasons behind errors (why they happen) are separate from the measurement of errors and pertain to quality assurance and improvement rather than to quality control. A process that demands multiple reviews will certainly tend to produce more accuracy than one that does not, but in the end it will prove too costly to be satisfactory, while in a quantitative vision efficiency is pivotal and is expressed as a relationship between the outcome and the resources to achieve it. In other words, resources must be proportioned to goals.

In an academic perspective, a correct translation is a translation with no errors; in a practice-oriented perspective, a correct translation is a translation where total errors are within the desired threshold in a quality index. Therefore one way to judge whether TQA on a project is complete is to measure translation defect density.

When dealing with TQA, a tool should be available to track any potential issues in a translation and guide the user in deciding whether or not these issues are relevant, and whether or not corrective action is necessary. For any TQA tool to work, explicit – and reliable – assessment criteria are required together with sampling rules for the extraction of representative samples for cases where the entire text is unsuitable for a comprehensive quality control due to size and/or complexity.

2.3. Sampling

Sampling is a statistical procedure for accepting or rejecting a batch of merchandise or documents through the determination of the maximum number of defects discovered in a sample before the entire batch is rejected. For an object to be measurable, it needs to be broken down into 'lots', standardized in terms of size and scope, and large enough to estimate and define limits both for the number and significance of defects found. Statistical sampling can be used to determine acceptability provided that acceptability criteria for inspection by attributes are set. The ISO 2859 series of standards (ISO 1985, 1991, 1995, 1999a, 1999b) can be used here as a reference.

Acceptance sampling is an important field of statistical quality control originally applied by the U.S. armed forces for the testing of bullets during World War II. In acceptance sampling, a sample is picked at random from a lot and, on the basis of the information yielded, a decision is made either to accept or reject the lot. Acceptance sampling is the middle-of-the-road approach between no inspection and 100% inspection. Its main purpose is to decide whether the lot is acceptable, not to estimate its quality, and it should be employed when:

- 100% inspection is too costly or takes too long;
- time or technology limitations are constraints;
- lot sizes are very large and the probability of inspection errors is high;
- the supplier's quality history is good enough to justify less than 100% inspection;
- potential liability risks are high enough to warrant some form of continuous monitoring.

For acceptance sampling to be effective, a lot acceptance sampling plan (LASP) must be implemented indicating the conditions for acceptance or rejection of the lot that is being inspected. These parameters are usually the number of different bad items (i.e. defectives) in a sample, and should vary in quantity and severity in direct relation to the importance of the characteristics inspected.

Average Outgoing Quality (AOQ) procedures are the best suited for small translation projects, since sampling is non-destructive, lots are 100% inspected and all defectives in the rejected lots are replaced with good units. In this case, all faulty lots are made perfect. The only defects, if any, left are those in lots that were accepted without (further) inspection. AOQ expresses the average nonconforming fraction that is shipped to customers – bad items are discarded but are not replaced with good ones:

$$AOQ(p) = \frac{(N-n)pP_A}{(N-n)pP_A + (1-p)N}$$

where P_A is the probability of accepting the lot, $(N-n)pP_A$ is the number of pieces that are shipped without inspection, and p is the non-conforming fraction. The numerator is the number of bad pieces that are shipped, and the denominator is the total number of pieces shipped.

To make assessment criteria, methods and tools unambiguous, AQLs (Acceptance Quality Levels) can be used allowing for tolerance and deviations (errors). AQLs should be agreed upon in a SLA and should specify the maximal percentage of non-conforming items to be considered as a satisfying process mean. Different AQLs may be designated for different types of defects. Usually,

an AQL of 1% is used for major defects, and 2.5% for minor defects. An implication of acceptance sampling is that a lot exceeding a given percentage of deviations from the AQL is unsatisfactory and must be rejected. At the same time, a high defect level (Lot Tolerance Percent Defective, LTPD) that is unacceptable to the consumer will have to be designated. AQLs imply that a level of non-quality exists in a product, and that defects can remain that will ruin a particular batch, despite being “acceptable” in general. This level represents a compromise between quality, quantity and price negotiated, even when – as is the case of translation – supply exceeds demand and the customer should be able to receive a flawless (no-defect) product.

To set AQLs, a simple defect prediction technique can be implemented to separate the defects found in a translation sample into two groups. Depending on the number of defects found in either of the two groups – but not in both – the defects that have not been found in the sample can then be estimated. This number gives approximately the number of defects in the entire project.

The Canadian federal government’s Translation Bureau has developed a complex system (SICAL, *Système canadien d’appréciation de la qualité linguistique*, Canadian Language Quality Assessment System), to assess 400-word chunks of translations from contractors. SICAL is based on sampling and a grading scale from A (superior) to D, depending on the number of major and minor errors. The Bureau’s goal is to deliver translations at levels A and B of the SICAL standard. In the Translation Bureau’s model, TQA is not confined to selecting a translator based on an analysis of sample translations to evaluate the translator’s skills; TQA is not a once-for-all task, *una tantum*, but it is a routine that is part of the production process. SICAL surreptitiously allows the Translation Bureau to decide whether to penalize contractors financially, thus partially recovering from costs through varying the remuneration according to pre-defined AQLs: a lower AQL gets a lower fee.

To calibrate a translation quality measurement tool or process, defects (errors) can deliberately be seeded in a translation to be checked. The ratio of the seeded defects found to the total number of defects seeded provides a rough estimate of the total number of translation defects yet to be found. It will then be possible to estimate what percentage of errors has been missed, and the variance in assessing the errors measured.

Among the many erroneous assumptions on quality, control uncertainty is probably the most impeding. A certain degree of ambiguity is obvious if assessment goals and criteria are not explicit and objective. This is why in the language industry, quality control is often confused with quality assurance to embrace editing. But much more time and money is spent on quality assurance than on the translation itself. In addition, a fully-fledged quality assurance process cannot do without inspections and auditing, as quality is not the result of

assessment and control procedures. These can lead only to the removal of defective products. Quality is a derivative property. In this perspective, it is not that hard to produce exactly what is requested when the assessment criteria are known.

3. Quality standards

The idea that quality can only be assessed against a set of specifications and requirements was introduced with the ISO 9000 quality standards. Since then quality has meant 'suitability for a purpose'. A quality system should be designed to specify expected and achievable quality levels, and be capable of generating a set of reports to detect deviations from a predetermined model.

Quality standards generally pertain to processes, to allow the customers of a certified company to receive the required goods or services in accordance with the agreed terms. Therefore requirements are pivotal for measuring quality after specific auditing, testing, and inspections on distinctive and standardized samples. Unfortunately, translation is rarely taught, and indeed thought of, as a repetitive and reproducible process, thus making auditing or inspection virtual tasks. Hence, to ensure quality, translation requirements must be both explicit and implicit. In the first case, quality level must be agreed with the customer on the basis of measurable parameters. The only measurable parameter in implicit requirements is suitability, corresponding to communication effectiveness, which is determined, in turn, by correctness and functionality.

3.1. The four rules of quality

In Peter Drucker's (1993) words

Quality in a product or service is not what the supplier puts in. It is what the customer gets out and is willing to pay for. A product is not quality because it is hard to make and costs a lot of money, as manufacturers typically believe. This is incompetence. Customers pay only for what is of use to them and gives them value. Nothing else constitutes quality.

Offering a better-than-acceptable level of quality without missing any deadlines, and at a reduced cost, requires considerable process innovation. Studies on evaluation techniques, standards to distinguish between minor and severe mistakes, and attempts to define what constitutes a good-quality translation have been argued over by many scholars and industry professionals. Quality is the responsibility of everyone in the organization and not exclusively that of the quality department – and quality improvement, contrary to traditional

belief, has a cost-reducing effect. Doing it right the first time may require an initial investment, but the impact in the long term generates many advantages outside the limited framework of quality.

Quality systems hinge on four basic rules:

1. write down what you do;
2. do what you have written;
3. substantiate what you have done;
4. reflect on how to improve what you have done.

In this view, quality is an endless work cycle. A cycle where deliverables are analyzed, proposed, developed, and delivered, then once again analyzed and improved on. A cycle of constant listening, observing, and quantifying, which will be refined and improved, will produce products more responsive to the needs of the users while meeting the customer's expectations.

Therefore, quality must be planned into a project and managed over the life of the project. Ensuring quality means accounting for the time for reviews in the project plan. It means taking the time to assess the needs of the user and setting aside the time to meet and come to an agreement on how quality will be measured and who will measure it. For a quality system to work, processes must be established and described according to the principles and criteria of the standards. Clearly, though, this is also the main difficulty in implementing quality standards. Yet, in most cases, taking the path to certification leads to awareness of where the inefficiencies lie and, after appropriate adjustments, to considerable process improvements as requirements must be thoroughly defined and detailed at each stage, while the system must be set up to ensure meeting them.

3.2. Quality is money

Value can be defined as the benefit of an activity minus its cost. When both benefit and cost of translation can be expressed in monetary terms, a monetary value can be calculated. A cost figure obtained through careful benchmarking can be used with greater confidence than a rough estimate of time and materials; while as long as benchmark costs are not known, translation will continue to be regarded simply as an expense rather than an investment. Costs can be calculated only when tasks are consistent and repeatable, and can be used to show the value added by quality. Measuring value added by translation means measuring the total value returned minus that cost. This value can be measured by measuring the change in value (the dependent variable) caused by a change in quality (the independent variable).

To the user, the cost of poor quality is in the waste of time and effort caused by inaccurate or unusable translations; to the customer, it is in extra support

time and the immense cost of revising translations. Of course, the greatest added-value of good quality translations is in increased customer satisfaction and the ensuing sales that this is likely to bring, both from the customer and from others who hear about the translator's performance.

During the first international conference on specialized translation in Barcelona in March 2000, Salvador Aparicio i Paradell illustrated the following formula to calculate the real cost of a translation:

$$q = t \frac{1+e}{1-e} + r + a$$

where q = quotation, t = translation, e = error rate, r = revision and a = accessories.

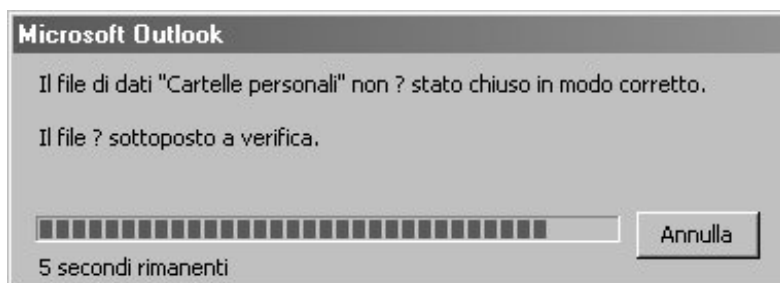
To guarantee quality standards, successful methods must be repeated and extended across projects, goals must be set, benchmarks must be established, records must be kept, and results must be assessed.

The value of effective communication is most frequently measured in the negative, that is, only if there are problems with effective communication, can figures be drawn that denote the extent of the problem. In the worst case, this negative example could be a lawsuit in which a client claims damages for several million euros or dollars because the handling of a machine according to the documentation has led to severe injury.

In localization, translation quality cannot be narrowed to linguistic properties (attributes). For example, in Windows XP the dial-up interface prompts the user with the following box "Verifying username and password..." (34 characters). In the Italian version, this became "Verifica della password e del nome utente in corso..." (53 characters, +18%), but the string which appears on the screen is truncated:



Again, when recovering a data file that has not been closed properly, the Outlook XP interface prompts the user with the following box:



In both cases, the translation is linguistically acceptable – even though a better translation than ‘rimanenti’ could have been found for ‘left’ – but the final screen vision impairs it. Not surprisingly, a common element in the diffidence of the general public towards open source software (OSS) is reportedly localization quality as performed by amateurs rather than by professional translators, making OSS only interesting for computer geeks.

Curiously, the following message box does not seem to have any negative effects on users, as ‘incorretta’ (instead of “scorretta”) is perceived as a minor inaccuracy and is skipped over, possibly with a smirk in consideration of the money spent on a product that was supposed to be of superior quality.



3.3. Customer satisfaction

Finally, customer satisfaction is the other side of the coin. It is the engine of, and the drive for, quality. Customer satisfaction can be measured from the customer's or in the service provider's perspective. From the customer's perspective, the levels of reaction to unsatisfactory service are the following:

1. disappointment: customers do not get what they really wanted;
2. allowance: customers accept a product whose quality is lower than expected;
3. trade-off: customers adjust their expectations;
4. settlement: customers' needs are met, but desires are not;
5. tuning: customers change their behavior to match the offering.

From the service provider's perspective, there is an equivalent scale:

1. fulfilment: the provider meets customers' expectations by giving them what they asked for;
2. satisfaction: customers' expectations have been met;
3. efficiency: typical offering has met customers' expectations;
4. equalization: operating efficiency is improved by leveling offering;
5. massification: customers are trained to ask for what is offered.

Quality is always listed as the highest priority; more important than deadlines, cost, and customer service. Nevertheless, trustworthiness is fundamental as most customers typically use only one vendor – and little time or money is allocated to translation itself.

For customer satisfaction to be measured, the relevant attributes must be determined including confidence, courtesy, friendliness, responsiveness, complaint handling and reliability. To trace customer satisfaction, a regular survey is necessary; one that can provide a statistical measurement of inbound and outbound deviations from the negotiated service level reported by customers. Tracked over time this will become a reliable quality index, and can be associated with in-process metrics to measure the effectiveness of reviews and the process over time.

Yet translation is an intangible service, circumstances regarding production are always different, and many factors can negatively impact on customer satisfaction. Consequently, customer survey results regarding particular situations can undermine a vendor's general perceived effectiveness. Therefore, it is necessary to guard against excess: the effort put into achieving customer satisfaction is sometimes extreme, even counterproductive. This will happen, for example, when the expectations ascribed to the customer have not been confirmed by the analysis. In these cases, there is a major risk of focusing on issues that the customer may have been unaware of – and are, in fact, immaterial – while leaving real issues unresolved and actual expectations unsatisfied.

The Canadian federal government's Translation Bureau admits that the ultimate test of the quality of a translation is client satisfaction. To measure client satisfaction and quality of translations the Translation Bureau implemented a Continuous Evaluation System based on sampling and periodic surveys. A survey, by its nature, cannot measure the emotional feeling toward an intangible service, and even with a standard set of rules, judgments will differ as interviewees will naturally base their feelings on different projects, done by different teams in different locations under different conditions. This also means that the wider the sample base, the more inconsistent the results of the survey will be, clashing with the fundamentals of statistics, a science where accuracy relates largely to the size of the sample; the smaller the sample size, the greater the bias. In addition, a deeply unhappy client who is not in a long-term

relationship with a vendor that s/he is determined to preserve will find it uneconomic to report any dissatisfaction and will more likely simply choose a new vendor. Also, customers tend to remember, and report, only major problems, which weigh heavily on overall satisfaction. Finally, as Jeffrey Gitomer (1998), the sales guru, put it,

Boasting about a near-perfect customer-satisfaction rating of 97.5 percent is a major mistake. That means 2.5 percent of your customers are mad, and they're telling everyone. And 97.5 percent of your customers will shop anywhere the next time they go to market for your product or service.

When running a customer satisfaction survey, even though virtually all customers are satisfied, they can still go for a competitor whom they also find satisfactory. Therefore, in creating customer satisfaction surveys, questions should be asked about expectations along with satisfaction. Measurements could then help predict the quality of the final completed product before actual completion. In-process metrics should then be developed by watching trends over time and correlating these trends with final quality.

This process of continuous improvement is called *kaizen* from the Japanese management concept for incremental adjustments introduced by Taiichi Ohno who was the assembly manager for Toyota in the 1940s and early 1950s, and developed many improvements that eventually became the Toyota Production System (Ohno 1988).

3.4. *Kaizen*

The *kaizen* method of continuous incremental improvements is based on traditional Japanese philosophy, assuming that every aspect of our life deserves to be constantly improved. *Kaizen* literally means change (*kai*) to become good (*zen*). When applied to the workplace, *kaizen* means continuing improvement involving everyone in an organization working together to make improvements 'without large capital investments'. The focus is on eliminating waste in all the systems and processes of an organization.

The key elements in the *kaizen* strategy are the willingness to change, a never-ending effort to improve and to communicate more effectively. Quality improvement and cost reduction are, in fact, compatible since quality is the responsibility of everyone in the organization and not exclusively that of the quality department. This means that everyone involved in a project should monitor quality at every stage of the process. Organized *kaizen* activities lead to the TQC (Total Quality Management) approach for improving performance.

Incremental improvements have a cost-reducing effect. The long-term impact of the doing-it-right-first-time philosophy generates many advantages

outside the limited framework of quality, making the initial investment worthwhile.

Traditionally, in order to verify the quality of a translation, a revision by a second translator is carried out, a practice that is certainly costly and time-consuming, especially because this work has traditionally been performed by senior translators. Eliminating most of the repetitive, measurable and predictable (formal) mistakes in advance would considerably reduce the time required for proofreading and correction work later, and what is measurable is also traceable. The clever project manager's motto is "deliver quality on time and within budget", a goal that can be achieved only through a combination of people, process, and technology.

4. The teacher's role

Students should be taught to devise and implement an overall project strategy. A project strategy makes translation requirements easier to collect and understand and apparent, even when they are not. A lack of standards, numbers, or ratios of quality allows ambiguity to arise since students, as future translators, are expected to deliver quality from the start, but will hardly find anybody capable of defining it. In fact, in translation classes educational goals are explicit, but how they will be pursued and monitored is left unsaid. Teachers are then called to play the unpleasant role of editors or reviewers who, no matter how necessary their corrections may be, send a demotivating message to translators: "You write poorly" or even "I write better than you do."

All translators eventually must confront editors or reviewers, but are rarely taught to view them as part of a collaborative endeavour to improve their work. To do this, students must be taught to work in a team and contrast the typical translator's disease, the self-referential attitude from isolation. This attitude sometimes leads editors – who possibly once were translators – to have a deserved reputation for making changes purely to demonstrate their authority. Even more frustrating is that some people gladly leave quality issues to the translator's expertise during the project, but turn into fierce critics after delivery.

Therefore, teachers should persuade students always to seek out an editor's assistance to make their lives easier once they become translators. In addition, by making error spotting, assessment and editing criteria explicit, the teacher can help students reduce subjectivity in judgment and learn how to develop their own metrics when reviewing or editing a translation. Also, telling a student what the teacher expects from them corresponds to clarifying requirements and making metrics explicit, thus making assessment transparent.

Finally, translation courses generally lack an 'economic' approach, with the associated investigation of the cost of errors, thus eluding the problem of

translation sustainability, which must be valid for the customer as well as the translator. It is to some extent equivalent to allocation efficiency by bringing about the best outcome for all by deriving the largest possible utility from any given set of resources.

Pricing strategies are crucial in this respect, as different requirements and jobs with different AQLs call for different offers. Also, tools that reduce source content to reduce translation costs are increasingly widespread. Therefore students must be taught to take full advantage of the appropriate technology to improve efficiency, use of resources, costs, and guarantee economic sustainability by standardization and large scale use, reliability, and affordability.

5. The role of the market

In most cases translation is not part of the customer's core business, who therefore considers it to be a non-critical purchase. Combined with the complexity of the supply market, increasing competition, and a more professional purchasing behavior, all this results in the perception of translation as a commodity. One indicator is the practice of auctioning for the assignment of translation projects. Translation is often at the end of a supply chain where all parties assume the preceding ones have performed their task to the best of their abilities. Therefore, even when translation is not considered as a commodity, it is sidelined, and is simply expected to be there.

On the other hand, according to the first law of socio-economics, in a hierarchical system the rate of pay for a given task increases in inverse ratio to the unpleasantness and difficulty of the task. As a result, vendor selection is usually based on generic business benchmarks rather than on the specific skills required to handle the translation process. In an ideal market, suppliers do not control markets, customers do, and no customer is willing to pay for poor-quality products. As customer power (pressure) is almost always exerted on prices, a lack in quality becomes a tangible element that must be taken into account when calculating a company's profit margin. Clients are interested in getting the lowest possible price while retaining the best service providers; conversely, vendors are motivated to get a fair price for their services and to resist price pressure.

Since auctioning is aimed at driving the price down, stressing quality makes no sense when the main parameter is the discount floor. The acceptable price is the one that is balanced to the ability to provide a service, rather than to the value of the auctioned item. According to the mathematician and Nobel laureate John Nash's theory of non-cooperative equilibrium, presuming that market players on average estimate the value of an item and their bids correctly, the

winning bid produces lower than feasible or even negative profit (Nash 2001). Winning against a number of rivals following similar bidding strategies implies that the winner's estimate is an overestimate of the item's value or an underestimate of a feasible contract bid conditional to the win.

Nobody is interested in driving the price below a sustainability level because this could put the supplier's reliability and capability to invest at risk, while adding considerable extra costs and risks on the customer's side. The extra costs involved in working with cheaper suppliers are for additional monitoring, while the extra risks concern rework and delays.

Only lower quality can be bought at a lower price; this is an old and well-known adage. A lower price often means that the supplier overestimated his/her capabilities, is probably working below his/her sustainability level and therefore has no reliable reserve or guarantee. This means that the quality could easily be lower than expected. *Pecunia non olet*, but translators are so focused on their 'art' that they pretend to ignore this saying or, worse, forget it, even though, when translation is not just a second-job option, money is and must be a real priority.

In short, quality must be proportioned to profit, and translators should be taught to think of their job in terms of making a living, and not as a form of art, and thus priceless per se.

In 2002, an Allied Business Intelligence research (ABI 2002) produced forecasts for a growth of the global translation market from \$ 13 billion in 2000 to approximately \$22.7 at the end of 2005. According to the same estimates, the publishing industry covers less than 5% of the market. Simply stated, and regardless of the roughly 40% cut in rates, this means that literary translation does not pay. Therefore, all estimates point to a market where economic sustainability is in the interest of both parties: the translation buyer and the translator provider. In literary translation, 'poetic' attributes prevail on functional features, thus exposing any assessment to severe subjective interference. Also, in literary translation, 'historicization' is crucial, and is the only acceptable filter for assessment.

Theorizing is a license to elude the questions discussed so far, and allows those involved to renege on verifications, thus justifying the otherwise factious dichotomy between 'practitioners' and 'theorists'. Yet, the first need the latter for the development of reference models, while the theorists need the practitioners to verify the correctness and relevance of their models. And since quality is a shared effort, any impediments in the quest for quality should be tackled in partnership with the customer. This will also increase the customer's understanding of why translations are priced differently for different types of manuals, and not only.

In sum, translators should learn to speak their customer's language, to explain how their services are different: in quality.

References

- ABI (2002) *Language Translation, Localization and Globalization: World Market Forecasts, Industry Drivers and eSolutions*, Allied Business Intelligence Research, Report code RR-TRAN.
- Drucker P.F. (1993) *Innovation and Entrepreneurship*, New York, Harper & Row.
- Gitomer J. (1998) *Customer Satisfaction is Worthless, Customer Loyalty is Priceless: How to Make Customers Love You, Keep Them Coming back and Tell Everyone They Know*, Austin, Texas, Bard Press.
- ISO (1985) *ISO 2859-2:1985 Sampling procedures for inspection by attributes – Part 2: Sampling plans indexed by limiting quality (LQ) for isolated lot inspection*, International Organization for Standardization.
- ISO (1991) *ISO 2859-3:1991 Sampling procedures for inspection by attributes – Part 3: Skip-lot sampling procedures*, International Organization for Standardization.
- ISO (1994) *ISO 8402:1994 Quality management and quality assurance – Vocabulary*, International Organization for Standardization
- ISO (1995) *ISO 2859-0:1995 Sampling procedures for inspection by attributes – Part 0: Introduction to the ISO 2859 attribute sampling system*, International Organization for Standardization.
- ISO (1999a) *ISO 2859-1:1999 Sampling procedures for inspection by attributes – Part 1: Sampling schemes indexed by acceptance quality limit (AQL) for lot-by-lot inspection*, International Organization for Standardization.
- ISO (1999b) *ISO 2859-4:1999 Sampling procedures for inspection by attributes – Part 4: Procedures for assessment of stated quality levels*, International Organization for Standardization.
- ISO (2005) *ISO 9000:2005 Quality management systems - Fundamentals and vocabulary*, International Organization for Standardization
- Nash J. (2001) *The Essential John Nash*, Princeton/Oxford, Princeton University Press.
- Ohno T. (1988) *Toyota Production System. Beyond Large-Scale Production*, Portland, Productivity Press.
- SAE (2005) *SAE Standard J24250 - Translation Quality Metric*, The Society of Automotive Engineers, SAE International.

Further Reading

- Al-Qinai J. (2000) "Translation quality assessment: strategies, parameters, and procedures", *Meta* XLV:3, pp. 497-519.
- Bonthrone R. (1996) "Screams in the quality jungle", *LISA Newsletter*, V:3, <http://www.fb-partners.com/pdfs/02-Screams-0996.pdf>, consultato il 2-3-2006.
- Brunette L. (2000) "Towards a terminology for translation quality assessment: a comparison of TQA practices", *The Translator* 6:2, pp. 169-182.
- Eckersley H. (2002) "Systems for evaluating translation quality", *Multilingual Computing & Technology*, # 47, 13:3.
- Hönig H. (1998) "Positions, power and practice: functionalist approaches and translation quality assessment", in *Translation and Quality*. Ed. by C. Schäffner, Clevedon / Philadelphia / Toronto / Sydney / Johannesburg, Multilingual Matters, pp. 6-34.
- House J. (1997) *Translation Quality Assessment: A Model Revisited*, Tübingen, Gunter Narr.
- House J. (1998) "Quality of translation", in *Routledge Encyclopedia of Translation Studies*. Ed. by M.Baker, London/New York, Routledge, pp. 197-200.
- Larose R. (1998) "Méthodologie de l'évaluation des traductions", *Meta* XLIII:2, pp. 163-184.
- Lauscher S. (1998) "Concepts of translation quality and quality assessment", in *Proceedings of the 39th Annual Conference of the American Translators Association*, Alexandria, VA, A.T.A. (CD Rom).
- Ling Koo S. & Kinds H. (2000) "A quality-assurance model for language projects", in *Translating into Success: Cutting-Edge Strategies for Going Multilingual in a Global Age*. Ed. by R. Sprung and S. Jaroniec, Amsterdam/Philadelphia, John Benjamins, pp. 147-157.
- Picken C. (ed.) (1994) *ITI Conference 7 Proceedings: Quality-Assurance, Management and Control*, London, Institute of Translators and Interpreters.
- Reiss K. (2000) *Translation Criticism - The Potential & Limitations. Categories and Criteria for Translation Quality Assessment*. Transl. by E.F. Rhodes, Manchester/New York, St. Jerome/American Bible Society.
- Schäffner C. (ed.) (1998) *Translation and Quality*, Clevedon/Philadelphia/Toronto/Sydney/Johannesburg, Multilingual Matters.
- Schiaffino R. & Zearo F. (2005) "Translation quality measurement in practice", in *Proceedings of the 46th Annual Conference of the American Translators Association*

<http://www.aliquantum.biz/Downloads/Translation%20Quality%20Measurement2.ppt>, consultato il 21-2-2006.

Williams M. (2004) *Translation Quality Assessment: An Argumentation-Centred Approach*, Ottawa, University of Ottawa Press.