



Defining Translation Quality¹

Geoffrey S. Koby (KSU)
Paul Fields (LinguaTech)
Daryl Hague (BYU)
Arle Lommel (independent scholar)
Alan Melby (LTAC)



ABSTRACT

Many problems in the translation industry revolve around customer dissatisfaction with someone's translation quality and disputes between translators when one is revising the work of the other. To determine whether someone has attained translation quality, one must be able to measure it. To measure translation quality, one must be able to define it. And to define translation quality, one must be able to define both *translation* and *quality*. Our two prior articles have addressed these terms respectively. Those articles do not resolve disagreements about these terms; indeed, we authors have unresolved disagreements among ourselves. However, the articles outline how different definitions can radically affect people's expectations about *translation* and *quality*. This third article builds on the first two and presents two contrasting definitions of translation quality: a narrow one and a broad one. The narrow definition matches with the narrow definition of translation in the first article and an emphasis on the transcendent view of quality in the second article, while the broad definition matches with the broad definition of translation in the first article and an integration of the five approaches in quality management from the second article. The authors strongly disagree about which definition of translation quality is better for the translation industry. Rather than pretending these articles are the end of this discussion, the authors hope to encourage a broad-based continuation of the debate that has taken place among the authors.

Keywords: Translation quality, Specifications, Metrics, Error categories, Broad, Narrow

RESUM (Tot definint qualitat en traducció)

Molts dels problemes de la indústria de la traducció giren al voltant de la insatisfacció del client amb la qualitat d'alguna traducció, així com amb conflictes entre traductors quan un revisa el treball d'un altre. Per determinar si algú ha arribat a la qualitat de traducció pactada, un ha de poder mesurar-la. Per poder mesurar la qualitat de la traducció, un ha de poder definir-la. I per definir la qualitat de la traducció, un ha de poder definir dos conceptes, traducció i qualitat. Els nostres dos articles anteriors han abordat aquests dos conceptes respectivament. Aquests articles no resolen les discrepàncies entre nosaltres. No obstant això, els articles exposen com les diferents definicions poden afectar radicalment les expectatives sobre traducció i qualitat. Aquest tercer article es basa en els dos articles anteriors i presenta dues definicions confrontades de qualitat en traducció: una més concreta i una altra més àmplia. La definició més concreta combina la definició

¹ This article is the first of trilogy of papers published in this issue: Melby *et al.*, Fields *et al.*, and Koby *et al.*



més estricta de traducció del primer article amb la visió transcendent de qualitat del segon article. En canvi, la definició més àmplia combina la definició àmplia de traducció del primer article amb una integració dels cinc enfocaments sobre gestió de la qualitat del segon article. Els autors estan totalment en desacord pel que fa quina de les definicions de qualitat de traducció és millor per a la indústria de la traducció. En lloc d'optar per tancar aquest debat amb aquests articles, els autors esperen haver donat peu a continuar amb el debat d'una manera més àmplia de com s'ha produït entre ells.

Paraules clau: qualitat en traducció, especificacions, mètriques, categories d'error, *narrow*, *broad*.

RESUMEN (*La definición de calidad en traducción*)

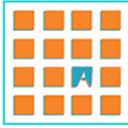
Muchos de los problemas de la industria de la traducción giran en torno a la insatisfacción del cliente con la calidad de alguna traducción, así como con conflictos entre traductores cuando uno revisa el trabajo de otro. Para determinar si alguien ha alcanzado la calidad de traducción pactada, uno debe poder medirla. Para poder medir la calidad de la traducción, uno debe poder definirla. Y para definir la calidad de la traducción, uno debe poder definir ambos conceptos, traducción y calidad. Nuestros dos artículos anteriores han abordado estos dos conceptos respectivamente. Dichos artículos no resuelven las discrepancias entre nosotros. No obstante, los artículos exponen cómo las diferentes definiciones pueden afectar radicalmente las expectativas sobre traducción y calidad. Este tercer artículo se basa en los dos artículos anteriores y presenta dos definiciones confrontadas de calidad en traducción: una más concreta y otra más amplia. La definición más concreta combina la definición más estricta de traducción del primer artículo con la visión transcendente de calidad del segundo artículo. En cambio, la definición más amplia combina la definición amplia de traducción del primer artículo con una integración de los cinco enfoques sobre gestión de la calidad del segundo artículo. Los autores están totalmente en desacuerdo en cuanto cuál de las definiciones de calidad de traducción es mejor para la industria de la traducción. En lugar de optar por zanjar este debate con estos artículos, los autores esperan haber dado pie a continuar con el debate de un modo más amplio de como se ha producido entre ellos.

Palabras clave: calidad en traducción, especificaciones, métricas, categorías de error, *narrow*, *broad*.

Introduction

With respect to translation, our first article reflects our consensus that definitions can be placed within a “landscape” defined by two axes: scope (narrow-to-broad) and specifications (absolute-to-relative). The “scope” axis asks how narrowly or broadly one defines the term *translation*. The answer to that question will determine whether a given activity, such as localization or transcreation, falls within translation’s purview. The “specifications” axis asks whether all of a requester’s requirements must be explicitly stated. An “absolute” stance asserts that some requirements are always understood and absolute and constant and can therefore remain unstated, while a “relative” stance asserts that in general, the best practice is to explicitly state all requirements as specifications, because they can vary from project to project. For both ends of the specifications axis, some degree of accuracy and fluency is required. The difference is whether the minimum levels depend on audience or purpose or both.

In our second article, we present multiple approaches to the question of how one defines quality. Based on Garvin (1984), we describe five approaches to quality (transcendent, product, user, production, and value), and we briefly present the contributions of several people who have shaped the discipline of quality management, according to which quality is what the customer perceives relative to expectations combining all five approaches. The second article concludes by addressing the issue of quality management’s relevance to translation. The authors disagree, for example, on whether raw or lightly post-edited machine translation that



meets the specifications for the project (that could probably not have been met using professional human translation) can be called a quality translation, even though it accomplishes the stated purpose and meets the expectations of the audience. This disagreement is related to two uses of the term “quality,” one in the discipline of quality management and the other in general usage, which tends toward a transcendent perspective. More fundamentally, the authors disagree sharply on whether the principles of quality management should be adopted by the translation industry.

In the present article, we build on the prior two articles to address *translation quality* and the ramifications of how people define that term. In doing so, we first list some points on which we generally agree. We then present two contrasting definitions of translation quality—narrow and broad—followed by a point-by-point analysis demonstrating how people supporting these opposing definitions evaluate significant translation-quality issues. After this analysis, we discuss the question of whether specifications need to be explicit and mention a few possible areas of research. We conclude with the most basic point of agreement among the authors: that they agree to strongly disagree with each other without walking away from the debate and to encourage a wider participation in this important debate about the nature of translation quality, with an eye to developing useful translation quality metrics.

Some Points of Agreement with Respect to Translation Quality

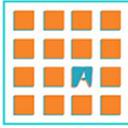
While we have disagreements among ourselves with respect to defining translation quality, we do agree on several issues. The first concerns how the language industry defines “customers.” This term includes both the party requesting the translation work and any parties who utilize the work product—the end-users. Thus, one can distinguish between these two types of customers as the *requester* and the *end-user*. The person or organization producing a translation for the customer(s) is the *provider*. Thus, we have agreed to use the terms *requester*, *provider*, and *end-user* for clarity.

Our second point of agreement concerns any product that can be understood to qualify as a translation. All translations carry an expectation of some degree of accuracy and fluency. Accuracy is a bilingual notion referring to the correspondence between the source and target text, while fluency is a monolingual notion referring to properties of the target text such as grammar, spelling, and cohesion.

Our third point of agreement pertains to the possibility of “perfect” accuracy and fluency. These two characteristics are always in tension because languages are not mechanical re-encodings of each other. Some slippage is inevitable, especially when sources move beyond banal propositional content (“the cat sat on the mat”) to anything where cultural expectations are an important factor. Perfection is therefore not a reasonable goal. If one rejects the notion of perfect accuracy and fluency, one can still propose a goal of “maximal” fluency and accuracy. This goal may seem fairly clear in isolation, but there are cases where a less-than-elegant target solution is best because a requester prioritizes fidelity to the source formulation (e.g., foreignized translation). Similarly, where the source text is defective in some fashion (e.g., it was produced by a second-language learner), a tension is created between accuracy and fluency. Should some infelicities in the source text be reflected in the target text, or should the target text improve on the source text and hide its flaws?

Our fourth point of agreement pertains to providers’ responsibility to requesters and end-users (the audience). Providers should endeavor to understand requesters’ purpose and take audience and purpose into account during the translation process. This emphasis on respecting a translation’s purpose has a long tradition in translation studies, beginning with Skopos theory (associated most closely with Hans Vermeer) and continuing with Functionalism (associated most closely with Christiane Nord).

Our last point of agreement concerns our conviction that both the language industry and translation studies urgently need a method to measure translation quality as objectively as



possible. That method should emphasize identifying problems that can be corrected. Any effort to measure translation quality is doomed to confusion without an explicit definition of translation quality.

Two Definitions of Translation Quality: Broad and Narrow

Given our own lack of consensus on defining translation in the first article and defining quality in the second article, readers should not be surprised that we have not reached consensus on a definition of translation quality. Below we present two contrasting definitions of translation quality, along with arguments in favor of one or the other. We label these definitions “broad” and “narrow.”

A Broad Definition of Translation Quality

Concerning the “scope” axis described above, a broad definition of translation categorizes many activities as “translation.” These include full translation (that is, segment by segment), summary translation, localization, transcreation, and even gisting (typically a rough translation done by a human or a machine translation system to help people decide whether a professional translation is needed). As to the “specifications” axis, this broad view asserts that no absolute specifications apply to all translation projects. Specifications should therefore be made explicit whenever feasible, relative to the audience and purpose. This broad view supports the following definition of translation quality:

“A quality translation demonstrates accuracy and fluency required for the audience and purpose and complies with all other specifications negotiated between the requester and provider, taking into account end-user needs.”

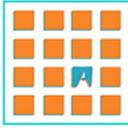
This broad definition fits with approaches to quality described in the second article. First, because the definition demands that all translations share innate characteristics—namely, certain degrees of accuracy and fluency—it acknowledges a “transcendent” approach to quality, as described by Garvin and the second article in this series. Second, because accuracy and fluency are characteristics that can be measured, this broad definition fits Garvin’s product-based approach as well. Third, because accuracy and fluency must reflect levels “required” by audience and purpose, the definition matches Juran’s notion that a quality product must fit its purpose. Fourth, the definition includes compliance with specifications, an emphasis found in both Crosby and Garvin (the “production approach”, which Garvin calls the manufacturing approach). Fifth, the definition acknowledges the end-user, recalling Garvin’s user-based approach and Feigenbaum’s emphasis.

Although the definition does not explicitly address Garvin’s value-based perspective, it acknowledges the need for negotiation between requester and provider on all aspects of the perceived value of the translation service, including target content requirements, delivery date, and cost. These aspects are part of structured translation specifications, as described in the appendix to the first article.

A Narrow Definition of Translation Quality

With respect to the scope axis described above, a narrow definition of translation views translation as text-centric. Other activities common to the language industry, such as summarization, transcreation, gisting, and certain aspects of localization, do not qualify as “translation” under this view. In addition, when machine translation is involved, this view requires a sufficient level of post-editing to ensure minimal levels of accuracy and fluency. Concerning the specifications axis, a narrow view posits that many requirements are absolute, i.e., subject to invariable minimum levels, although there are many variables that must be specified (i.e., the traditional translation brief). The foregoing assumptions support a narrow definition of translation quality as follows:

“A high-quality translation is one in which the message embodied in the source text is transferred completely into the target text, including denotation, connotation, nuance, and style, and the target text is written in the target language using correct grammar and word



order, to produce a culturally appropriate text that, in most cases, reads as if originally written by a native speaker of the target language for readers in the target culture.”

This narrow definition, from a quality management perspective, could be viewed as a special case of the broad definition, with various specifications predetermined. Projects outside the scope of the narrow definition of translation are likewise outside the scope of the narrow definition of translation quality. Thus, these two definitions of translation quality differ in their range of practical applications. From the perspective of a narrow view of the definition of translation, the broad definition of translation quality is too broad and should be viewed as covering translation quality management rather than just translation quality. The choice of one definition of translation quality over the other will influence the framework for developing translation quality metrics and have practical consequences for measuring translation quality (TQ). The rest of this article presents additional TQ issues in a debate format outlining how a broad or narrow definition shapes one's answer.

Some Additional Points of Debate

Should specifications always be explicitly stated?

Both the translation industry and translation studies urgently need a way to compare different sorts of translation as objectively as possible, with an emphasis on identifying problems. The ultimate goal in all cases is improving accuracy (adequate transfer of meaning) and fluency (correct target-language usage). However, some hold that this goal must be achieved within the context of agreed specifications, while others hold that extensive specifications going beyond the traditional translation brief may not be necessary in all cases.

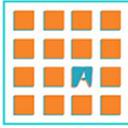
The broad TQ view implies that providers should work with the requester to understand the requester's requirements and the end-users' needs. Furthermore, those requirements and needs should be stated as formal specifications. Creating specifications is a discovery process that brings unstated requirements into focus. As such, there is no such thing as a “zero-requirements” work product.

In contrast, the narrow TQ view implies that explicit specifications are not always needed. For example, requesters often do not know what the specifications are or should be. Furthermore, from a narrow view of TQ, most translation assignments fall within the “normal” standards of professional human translation, for which explicit specifications going beyond a general description of the text type and possibly the target audience (i.e., legal topic, into US English) are unnecessary, as the text types are typically translated into largely equivalent text types in the target language. Finally, providers who seek to establish explicit specifications about every issue can irritate the requester without improving the resulting translation.

Do providers have an ethical responsibility to meet minimum service levels regardless of what the requester asks?

Many people affirm that often neither the end-user nor the requester knows what a particular project actually requires. Under such circumstances, providers have an ethical mandate to deliver a translation product that meets certain minimum standards that are absolute, that is, independent of the type of the audience and purpose. A narrow TQ view would support the existence of such an ethical mandate, while a broad TQ view would not. Instead, the broad view would encourage providers to negotiate requirements—perhaps even a set of default requirements—with the requester.

Monolingual requesters may initially focus on price and be willing to accept a “rough” translation. Unfortunately, these requesters frequently misunderstand just how rough a raw or lightly post-edited machine translation can be. Furthermore, they may not anticipate the comprehension challenges monolingual end-users face. In such cases, requesters often complain about the quality received, even though the translation complies with specifications. Given this situation, some people argue that even if a requester asks for something that does not



meet minimum standards, the translator still has an ethical responsibility to meet that minimum. Furthermore, others may argue that for *most* situations, providers would be *wrong* to knowingly submit something that does not meet a minimal level of accuracy and fluency. A contrasting view is that translators who deliver a higher level of accuracy and fluency than identified in the specifications are ethically obligated to inform the requester, especially if this variation involves additional cost to the requester.

Regardless of whether a requester approves a set of specifications where “style” is not considered important, requesters frequently change their minds when the final translation is delivered. In a world where the customer is always right, referring back to these specifications is cold comfort for the provider. From a practical perspective, therefore, providers have a strong incentive to make the translation as fluent and accurate as possible even if the requester claims not to care. In some cases, however, requesters may willingly accept low levels of accuracy and fluency if the provider first sends some sample pages and provides other clear communication to confirm shared expectations.

Is there a particular set of error categories that should be included in every TQ metric?

Both the narrow and broad views agree that people can use multiple error categories, including the broad categories of accuracy and fluency, and specific error categories within broader error categories, to create a TQ metric. Disagreement sets in when one attempts to list the error categories that are absolutely necessary. A narrow view of TQ holds that any translation quality metric must necessarily include at least the following error categories: (1) meaning transfer; (2) terminology; (3) domain-specific writing quality; (4) domain-independent target-language accuracy. In contrast, a broad view would simply argue that some situations—e.g., time-sensitive intelligence work—might not require all of these error categories. The parties can therefore negotiate these error categories as they desire.

Can stakeholders establish minimal acceptable service levels for general situations?

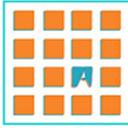
Perhaps one way to address quality concerns that would satisfy both broad and narrow TQ views would be defining explicit minimal acceptable service levels for general situations and making these levels known to both requesters and providers. Parties could then negotiate higher thresholds (e.g., for translating a nuclear disarmament treaty) or lower ones. But if lower thresholds are defined, the parties should do so consciously, and the request should be motivated. A help-desk translator’s job brief, for example (see second bullet point below for details), could read as follows: “We know that you are normally expected to meet certain requirements for fluency, but we want you to disregard them and instead focus only on making sure that certain keywords and relationships are represented accurately so that the database can locate the appropriate answers.” This brief could then acknowledge that minimal professional standards exist, but that there is a good reason to break them in this case. So perhaps quality evaluation standards could provide guidance for generally accepted quality levels and explicitly state that deviations must be agreed upon by all parties in advance.

How does one measure TQ when informed requesters deliberately accept less-than-maximal accuracy and fluency?

One can imagine a translation request in which specifications do not demand the translator’s best and instead require substantial deviations from an ideal situation. Examples are relatively easy to find in time-sensitive materials, such as the following:

- An intelligence analyst is producing translations of intelligence intercepts. The information is highly time-critical, and if the analyst takes time to revise and fix non-critical problems, the translation will quickly lose value. So, the specifications might actually state that any spelling, grammar problems, etc., that do not compromise the ability of the intended audience (e.g., military leaders) to make their decisions should NOT be fixed.

A second example is perhaps more to the point in the commercial world:



- MT is used to try to match user queries to an existing database of support items. If MT is not able to identify a valid match, then (and only then) human translators are looped in. The human translator (known as an “operator”) then has to decide whether to try to fix the MT output or retranslate it from scratch. If the operator chooses to retranslate, the retranslation’s ONLY short-run value is in identifying relevant items in the database. From a business perspective, any time spent on improving accuracy or fluency beyond what is needed to identify the relevant database content is wasted. Because this translation will never be used for an outside client, the text’s appearance does not matter at all. Therefore, the operator should NOT clean it up or polish it in any way.

From the industry client’s perspective, the operator represents a rather high-value position for a translator. By employing the operator, providers can significantly expand client support. Providers simply could not afford to do so with purely human translation, nor could they provide it at all with pure MT. The operator’s job is *not* a task for dummies, but requires great skill. Specifically, the operator provides an absolutely vital service that requires very fast turn-around and considerable understanding of products and services--and the MT system involved. But the translation output would not meet normal industry standards and in fact SHOULD NOT do so. Meeting normal standards for accuracy and fluency adds no value and in fact *subtracts* value by slowing the human operator down.

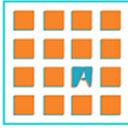
Proponents of a broad definition of translation quality point out that if stakeholders want to define quality in a way that reflects the scenarios in which translators may work (now and in the future), those scenarios must include varied requirements. Proponents of a broad definition therefore find the narrow definition of translation quality too rigid and not sufficiently comprehensive, since it is associated with a narrow definition of translation that categorizes summary translation, localization, and light post-editing, among other language industry activities, as secondary or additional activities that are separate from translation *per se*.

Proponents of the narrow definition of translation quality point out that a narrow definition is consistent with the intuition of many professional translators and reflects the way they think and work. A few years ago, there was a serious debate at the NATO translation office in Brussels. One of the authors was there. After several hours of discussion a veteran translator rejected a broad definition of translation quality very similar to the one given in this article, on the basis that a transcendent view of quality is the only possible view for a professional human translator. She and other proponents of the narrow definition of translation quality take the position that the broad definition of translation includes aspects that are separate from translation and can be performed by monolinguals.

Conclusion

Throughout this three-part series, we have addressed the challenges of defining *translation*, *quality*, and *translation quality*. We have found that “translation” can be located within a landscape defined by axes of scope and specifications. The authors have not reached consensus on whether a broad or narrow definition of translation is better for the entire range of stakeholders in the translation industry: requesters, providers, and end-users. Their only point of agreement in the first article is that it is important for any two people talking about translation quality to first identify where their respective definitions are “planted” in the landscape we describe, in order to avoid talking past each other.

Concerning the applicability of the discipline of quality management to the translation industry, the authors also failed to reach consensus. Their only points of agreement are (a) that it matters deeply whether the translation industry adopts or rejects all of quality management or emphasizes a transcendent view and (b) that a decision between emphasizing a quality management perspective and a transcendent perspective cannot be made independent of a choice regarding the definition of translation, including a discussion of whether there are absolute specifications that are independent of audience and purpose. In the present article, two



contrasting definitions of translation quality have been presented. Proponents of the broad definition see the narrow definition as a special case of the broad definition. Proponents on the narrow definition are not ready to accept this claim.

Concerning future research related to translation quality, statistical analyses are needed to show the breakdown of the various types of translation and peri-translation activities.² Additional research is needed on how the language industry (which itself must be carefully defined) perceives traditional text translation, localization, and other forms of adaptation and how the industry is evolving. How much of the language industry falls under the narrow definition of translation and how is market segmentation changing? Which of the 21 translation parameters in the appendix to the first article are considered part of translation quality, as a product and as a service, to which stakeholders? And how does relevance of parameters relate to one's choice of a broad or narrow definition of translation and translation quality?

Regardless of research on how people in the translation industry currently perceive translation quality, the authors agree that progress on systems to develop and apply translation quality metrics will be hindered unless everyone working on them lays out three definitions:

- Definition of translation (narrow, broad)
- Definition of quality (relative to Garvin's five approaches, embracing all or some)
- Definition of translation quality (narrow, broad)

The authors are not indifferent about which definitions are used. Some of us strongly believe that adopting the broad definition of translation quality and all that goes with it in quality management is very important to progress in the translation industry, in particular unifying the industry. However, others of us strongly believe that narrow definitions are sufficient.

Two things are certain: (1) translation quality metrics must be built on a well-defined foundation including at least clearly stated definitions of translation, quality, and translation quality; and (2) the important debate about which kind of definition to use, broad or narrow, is not over.

The authors fervently hope that the trio of articles concluding with this one will promote further meaningful and beneficial debate among language industry professionals.

References

- Crosby, Philip B. *Quality Is Free: The Art of Making Quality Certain*. New York: McGraw-Hill, 1979.
- Deming, W. Edwards. *Out of the Crisis*. Cambridge, MA: MIT Press, 1986.
- Feigenbaum, Armand V. *Total Quality Control*. New York: McGraw-Hill, 1983.
- Garvin, David A. "What Does Product Quality Really Mean?" *Sloan Management Review* 26.1 (1984): 25-43.
- Juran, Joseph M., and Joseph A. DeFeo. *Juran's Quality Handbook: The Complete Guide to Performance Excellence*. 6th ed. New York: McGraw-Hill, 2010.
- Nord, Christiane. *Translating as a Purposeful Activity: Functionalist Approaches Explained*. Manchester: St. Jerome, 1997.
- Vermeer, Hans. "Skopos and Commission in Translatorial Action." *The Translation Studies Reader*. Ed. Lawrence Venuti. 2nd ed. New York: Routledge, 2004. 227-238.

² One problem in this case is a lack of reliable statistics on the translation industry. It has been argued (by Durban, for instance) that much of the high-end or "premium" market work is performed by boutique translation agencies or individual translators who do not report statistics.