



FIT Research Task Force Position Paper

Translation Grades

FIT, the voice of associations of translators, interpreters and terminologists around the world, sees the need to state its position on translation grades. This position paper will discuss translation grades as defined in the CSA Research report “Making the (Translation) Grade: Tying Translation Quality to Stakeholder Requirements and Use Cases” (Lommel & DePalma, 2022), and some of the issues raised during the webinar organised jointly by CSA Research, a research company focused on the global content and language services markets, and FIT North America, held online on 9 March, 2023. We will also consider other articles on the subject¹ and, more broadly, the ways in which public, private and NGOs’ translation services are currently offered on the market². A caveat, as of date of sharing this latest version of the paper with FIT members (February, 2024), the concept of “grades” has already been included in 2023 international standard ASTM F2575-23e2: Standard Practice for Language Translation.

Introduction

With the growing demand for translations worldwide and the significant development of the language sector on a global scale, the question of translation quality standards has, once again, come to the fore. Standardisation organisations, such as ISO, have developed specific quality standards for translation services, protecting both the suppliers of translations and their clients. Recently, however, the rise of the use of machine translation (MT), and especially that of fully automated translation, which is broadly viewed as the solution for obtaining more or less satisfactory translations quickly and cheaply, together with the generalisation of the practice of post-editing (PE) (Gouadec, 2010: 273–275; Polikar 2023)³ seems to be challenging those standards. In other words, the standards written or revised before the boom of MT and post-editing in the translation sector might not fully account for these recent developments. This is why the question of defining translation categories has arisen (Way 2013), apparently to simplify the transaction between a language service provider (LSP), be it a company or an individual translator, and a client. Indeed, one must bear in mind that clients do not always understand the translation process: they simply want to get certain content in another language, or in

¹ Especially Zetzsche 2022 and Melby 2022.

² A detailed discussion of areas such as literary translation, audiovisual translation and transcreation is outside the scope of this paper.

³ PE, which starts with a translation produced by a machine, should not be confused with revision, which begins with a translation made by a human being. The different types of errors encountered require different cognitive processes, depending on the activity in which the individual is engaged: post-editing or revision. It is also necessary to distinguish between PE and human translation involving MT as an optional resource. Please see FIT Position Paper on Post-Editing (May 2021).

multiple languages. The various stakeholders may have a different understanding of how to define a translation, and what the translation process entails. In this sense, an agreement on translation categories could be a useful tool for the pre-production discussion between a requester and a provider of translation services to clarify and align expectations.

The translation market recognizes three types of machine translation: “light post-editing”, “full post-editing” and completely automated. By way of example, the Translation Centre for the Bodies of the EU⁴ has been offering light post-editing as of December 2021 as part of its added-value services. This new service aims to provide clients with machine-translated texts that have been revised lightly by one language professional (instead of two, as in the standard full post-editing service) so that the output is comprehensible. The service is offered when “the final text is not intended for publication and is mainly for information ‘gisting’, i.e. communicating the essential meaning without necessarily being perfect in the target language. The output will be a comprehensible text in which no important information has been accidentally omitted, and no incorrect information has been added”. They also offer a completely automated service for the translation of documents using a combination of the Centre’s translation memories and custom machine translation engines, without any quality control by translators or preparation of the source documents for technical processing or formatting. This service is recommended if the intention is to quickly grasp the overall message of a text written in a foreign language or have a translation for internal use only.

Other usages and types of requests can be found on the market, with the request of a “good enough translation” being one of them; the opposition between two types of purposes, that is, “for information purpose” and “for publication purpose”, is another one. These formulations are somewhat vague and, therefore, flawed.

Considering the above, would the requester find the new framework of translation grades more understandable and useful in the transaction with the producer of a translation product or service? Would translation grades facilitate the everyday life of translators and help maintain their professional status, including rates?

A need for translation grades?

The CSA report suggests that the purpose of translation grades is to offer a transparent way to determine what type of translation is suitable for a specific use. Specifically, the three translation grades proposed in the report –high, medium, and low-grade translation–⁵ differ from each other in how accurately the source content is conveyed and how fluent the target language is. The objective is “fitness for purpose”, i.e., translations that meet the clients’ requirements and expectations when the translation is used. It is easy to see, in the view of the authors, why such grades may be necessary. Clients are not always knowledgeable about translation, so they may not understand, for example, what raw machine translation can and cannot be used for, or what specific requirements exist for

⁴ See: Translation services | Translation Centre for the Bodies of the EU (<https://cdt.europa.eu/en>). Other examples can be driven from the OECD, which offers raw machine translation and post-edited machine translation services. The former is an NMT (Neural Machine Translation) based on OECD terminology, and establishes the result is meant “not for official use”; the latter is also an NMT based on OECD terminology, but features corrections made by OECD translators and establishes the result is “official use not recommended”.

⁵ Gouadec (2010) had a different formulation: “(1) rough-cut, (2) fit-for-delivery (but still requiring minor improvements or not yet fit for its broadcast medium), and (3) fit-for-broadcast translation (accurate, efficient, and ergonomic)”; he also suggested a fourth category, that of ‘fit-for-revision’, to describe translations that can be revised within a reasonable time at a reasonable cost.

specialised fields such as legal translations. Therefore, grades can help LSPs and their clients have a transparent conversation about needs and expectations. As the report points out, increasing information symmetry and reducing confusion in the translation process are important goals. However, the grades system, as presented in the report, also has some shortcomings. Some of the argumentation is flawed or missing important pieces, and the final model is therefore slightly unclear and not as helpful as it purports to be.

Translation as a product, a process, or a service

It is problematic to discuss translation solely as a product or commodity, comparable to bolts and eggs. As the authors of the CSA report themselves point out, “the translation industry straddles product and service categories” (page 9). This means that, in order to understand the nature of translation, we should consider its status as a service delivered by experts when designing the grades (or “categories”) and the criteria that determine these grades. Given that each text (and each context of use) is different, and because communication is dynamic by nature, translation cannot be assigned the same kind of stable technical criteria as bolts and eggs. It would be preferable to use examples that are more comparable to translations, and that allow us to consider the service element. Looking at the way other expert or professional services are characterised could provide useful comparisons that are not available when comparing translations to products and their characteristics. To resolve the dichotomy between service and product, we could say that translation, in the same way as law or accounting, is a service that results in a product, and belongs to the tertiary sector of society (Barabé 2021).

The CSA report states that grades are “categories assigned to products with the same broad functional use, but different technical requirements”. The idea of “technical requirements” is difficult to apply to translations. If a technical requirement for a bolt is its size or material, would the technical requirements for a translation be the number of words, the font, or perhaps the file format? To some extent, it would be possible to establish different types of translations by determining, for example, the format of delivery (e.g., whether they are just text files or formatted according to the client’s layout specifications), but that type of technical requirement is a rather minor aspect in terms of the translation itself. Furthermore, those technical criteria are not mentioned, so in fact the report overlooks the basic technical criteria and discusses the translated texts instead.

Are translation grades different from quality levels?

Focusing on the translated texts means that the grade descriptors are based on aspects of translation quality, as both correspondence and fluency are indeed quality criteria for a translated text. Yet, the report states that quality levels do not work as criteria for grades. This presents something of a discrepancy in how the grades have been determined: they are intended to be different from product quality considerations, but they use quality aspects as descriptors. The grades are admittedly different from the various quality assessment/assurance methods, but they rely on an understanding of what makes a ‘good’ translation. Such an approach unavoidably leads to the grades being seen as different levels of product quality, even if each level can be accepted as fit for a specific purpose.

What ISO standards state

It should also be noted that existing quality assurance methods already cover some of the concepts proposed in the grades system, and those processes have been designed to maximise objectivity and suitability to clients’ needs. For example, when quality is evaluated using quality metrics, the

translator's compliance with the clients' requirements is also checked, so the clients' specifications are already a factor in the assessment.

In fact, ISO 17100:2015/Amd 1:2017 Translation services — Requirements for Translation Services, establishes that clients and translation service providers may enter into agreements concerning the project specifications and implements a post-production step called "verification", i.e., a confirmation by the project manager that specifications have been met.

In ISO 17100, Annex B describes the "Agreements and Project Specifications". Some of the items contain references to "accuracy (including terminology) and fluency (e.g., syntax, spelling and lexical cohesion)"; whereas the CSA paper states that "correspondence" refers to a combination of accuracy and the correct use of terminology, and "fluency" is "the linguistic notion that content complies with all grammatical, spelling, stylistic, and typographical norms of a language". The definition of fluency seems to be rather similar, while that of accuracy is slightly different: In the ISO standard, "accuracy" encompasses the correct use of terminology while, in the CSA paper, another term is presented, that of correspondence, which has a broader scope and encompasses accuracy and terminology.

The examples of different degrees of correspondence and levels of fluency of the CSA paper listed on page 12 for marketing content (adaptation of the source/high fluency), technical manuals (precise correspondence/lower fluency), or legal filings (high correspondence and fluency), are not unique to grades. Such context and text-specific characteristics of the translation task are typically determined in a translation brief and reviewed after the project has been completed. In addition, it could be said that the way in which product quality is referenced in this report risks being more subjective than a carefully designed quality assessment framework, in which the reviewer's judgement is based on assessing an error category taken from an error matrix and allocating a severity score. Moreover, if the translator objects to it, they can justify their decision before the reviewer issues the final score. If they do not agree, a third-party reviewer is summoned to give their opinion. In the case of translation grades, no such process has been described, and the report does not offer a comprehensive definition of product quality. Thus, the basis for determining correspondence and fluency in the grades system is not entirely transparent.

The report states that the grades focus on aspects of the product of translation work and not, for example, on the translation process. This approach is contrary to translation quality management standards, where processes are certified (see ISO 17100 and other related standards). In addition, if we want to avoid thorny discussions on product quality, one way to do so would be to foreground the process instead. That approach would allow us to characterise appropriate uses for different kinds of translation, such as full professional translation, transcreation, sight translation, full MT post-editing and/or light MT post-editing⁶, and even to outline contexts in which the use of raw MT without any professional involvement may be possible.

In other words, process-based grades would explain what type of translation process constitutes a high level of service and when that would be an appropriate choice, and conversely, in which cases something more stripped down would suffice. That approach could be presented as a more objective and clearly defined way to align expectations, as it would be detached from the potentially subjective

⁶ ISO 18587:2017(en) "Translation services — Post-editing of machine translation output — Requirements" states: *full post-editing*: process (3.1.5) of post-editing (3.1.4) to obtain a product comparable to a product obtained by human translation (3.4.3)/ *light post-editing* (3.1.6): process of post-editing (3.1.4) to obtain a merely comprehensible text without any attempt to produce a product comparable to a product obtained by human translation (3.4.3).

discussion on what translation quality means. In essence, this approach would start with describing what process leads to a high grade, and then supplement this by explaining what type of translation each approach can be expected to produce. Such a description would not attempt to provide criteria for optimum quality. Instead, it would determine the circumstances under which optimum quality can be produced, and it would also describe use cases in which clients would need such optimum quality. Product quality could then be evaluated through established quality assessment procedures that would examine whether the expected quality has been reached. Ideally, criteria for both process and product quality would take into account the needs and preferences of the end users, i.e., the readers of the translation⁷. Alternatively, we could see a process-focused approach as risk-based, i.e. describing what kinds of risks are involved in different approaches to translation tasks, and what risk-mitigation efforts are available in various types of processes. A focus on risk would allow clients to consider when they might need the risk-mitigating services of LSPs and professional translators, and when they are prepared to assume the risk of using automated translation.

LSPs (Language Service Providers) and individual translators

A process-based approach would also remove the difficult assumption that professional translators may be expected to lower their standards to produce a medium or low-grade translation, or that LSPs might want to hire or retain “low-grade” translators for certain translation tasks. That is not how professionals or LSPs operate, so the concept of grades can be misleading. If the focus was on the process instead, the characteristics of professional translation work could be highlighted more easily. In fact, the contrast with process focus reveals another problematic aspect of this grade proposal: the absence of the individual professional translator as such.

On page 19 of the document, under the heading “How All Parties Can Use Grades”, only content creators, i.e., clients requesting translations, and LSPs are listed. Missing entirely is the individual translator, who is the one tasked with the translation itself. While the definition of LSP can include individual translators, it does not appear to do so in this context. It would be crucial to add the translator’s point of view and describe how these grades are to be incorporated into the work of professional translators, how translators benefit from the grades, and what the translator’s role is in negotiating the grades and associated specifications. Practitioners can suffer from the same kind of information asymmetry as clients, so in order to fully eliminate that problem, translators must be given due consideration. Again, a more process-based grades system would naturally involve practitioners, but even a product-based approach should be able to acknowledge them as crucial participants in translation projects.

Professional versus non-professional translation

Furthermore, it is important to make the position of the professional translator and the LSPs clear in the grades system. For example, on page 17, there is a reference to “community-managed translation”, a term used to indicate crowdsourcing, which is a controversial way of producing translations. The use of raw machine translation or any other method not involving professional translators may also raise questions. These translation methods are certainly used in various cases, and one benefit of a grades

⁷ See Suojanen et al. 2015 and Suokas 2019 for an example of how end users could be foregrounded in the translation process by applying the concept of usability.

system could be to clearly define the contexts or situations in which such approaches are sufficient⁸. However, it is also crucial to offer a clear distinction between the role of a professional translator or an LSP on the one hand, and the position of non-professional entities on the other hand, to define what the expectations are of a professional service. LSPs work with professional translators and appropriate translation technology, providing different types of services such as translation, post-editing and transcreation; they are always expected to adhere to professional standards and ethics, as stated in the codes of ethics issued by the translators' associations they belong to or by the LSPs themselves, thus minimising risk to their clients. Any proposal for translation grades or similar systems should not undermine the efforts of translators' associations to advance the professionalism of translators and to build a sustainable professional environment.

One often unspoken element of a professional service is trust: the client can trust that a professional LSP and a professional translator will deliver a high-quality service that serves the client well. Non-professional actors, on the other hand, exist largely outside this system and should be clearly distinguished from professionals in the description of grades. The grades system should state the unambiguous expectation that non-professional or fully automated options are those that companies or individuals can produce privately as part of their own processes and at their own risk, and therefore are mostly separate from the commercial translation market. It is also necessary to restrict the non-professional uses quite narrowly so that the grades do not give the appearance of advocating for expanding such practices in professional contexts.

In conclusion

It may be understandable that both clients and LSPs need to use a consistent, broadly implemented, and objective system of translation categories. Such a framework could also be useful for translators to enhance their work and protect their rights. Indeed, ensuring that the client understands what constitutes translation quality and process quality, and when the highest possible quality is needed, gives visibility to the expertise of professional translators. Especially now that machine translation and post-editing are so prevalent, and clients may not always understand what their translation needs are, transparent descriptions could protect all parties and allow for productive collaboration.

However, this model would benefit from additional work to fulfil that need. We should acknowledge the diverse nature of translation as an activity that includes a wide spectrum of practices, from unedited machine translation used by companies and individuals at their own risk, to a highly specialised expert service performed by professional translators. The framework needs to transparently identify the use cases and processes in which professional translation does not play a role and explain the risks this practice entails to users. It must also explicitly account for the types of cases where professional translation is crucial, and what the expectations are for these processes. It needs to recognise the position of professional translators in the different categories, clarify where professionals are not involved, and present the nature of professional translation as an expert service where ethics and trust play a central role. While some user needs may be served by automated or semi-automated processes where premium quality is not a priority, a framework of translation categories should particularly discourage a view of translation as a mass production activity and translations as mere commodities.

⁸ See Nurminen 2021 for a thorough discussion on how raw MT is being used in practice and on factors that influence the usefulness of raw MT as a communication tool.

As such, even though the concept of “grades” is already mentioned in some translation standards, we feel that the grades proposal would need additional research, and could evolve towards a service-oriented model that considers translation both as a process and as a product and takes into account the risks and risk mitigation elements associated with different kinds of translation practices. As a federation representing associations of translators, we remain open to discussing this issue further in the best interest of the entire profession.

References and further reading

Barabé, Donald. (2021). “Translation Status – A Professional Approach”, in *JoSTrans*, Issue 36. https://jostrans.soap2.ch/issue36/art_barabe.php

Barabé, Donald; Lommel, Arle; Elbaz, Pascale; Lank, Steve; DePalma, Donald. (2023). Categories of Translation, Translation Commons Social Channel, https://www.youtube.com/watch?v=2Ik_mTbLgq0

Gouadec, Daniel. (2010). Quality in translation. In: Yves Gambier & Luc van Doorslaer (eds), *Handbook of Translation Studies*, Volume 1. Philadelphia: John Benjamins, 270–275.

FIT’s Research Taskforce. January. (2022). *FIT Position Paper on Translation, Localisation and Transcreation*. PDP 202201 Translation Localisation Transcreation EN.pdf (fit-ift.org).

FIT’s Research Task Force. (2021). *FIT Position Paper on Post-Editing*. https://library.fit-ift.org/public/Publications/positionpapers/PDP_202105_Post_Editing_ENFR.pdf.

FIT’s Research Task Force. (2017). *FIT Position Paper on International Standards*. https://library.fit-ift.org/public/Publications/positionpapers/PDP_201706_International_Standards_ENFR.pdf.

FIT’s Research Task Force. August. (2021, updated). *FIT Position Paper on Machine Translation*. (August 2019, updated): https://library.fit-ift.org/public/Publications/positionpapers/PDP_201908_Machine_Translation_ENFR.pdf.

Lommel, Arle, DePalma, Donald. (2022). *Making the (Translation) Grade. Tying Translation Quality to Stakeholder Requirements and Use Cases*. CSA Research. https://insights.csa-research.com/reportaction/305013407/Marketing?_gl=1%2A4zc4cr%2A_ga%2ANDkyNDE1ODQ5LjE2NzkzMjY3NzU.%2A_ga_LL76TPHNBP%2AMTY4MTMzMTQ0My4yLjEuMTY4MTMzMTQ1OC4wLjAuMA

Melby, Alan. (2022). “Grades of Translation as a Marketing Strategy in the Era of AI”, *The Tool Box Journal*.

Nurminen, Mary. (2021). Investigating the Influence of Context in the Use and Reception of Raw Machine Translation. PhD Thesis. Tampere University Dissertations: 515. <https://urn.fi/URN:ISBN:978-952-03-2199-4>.

Suojanen, Tytti; Kaisa Koskinen & Tiina Tuominen. (2015). *User-Centered Translation*. Oxon: Routledge.

Suokas, Juho. (2019). User-Centered Translation and Action Research Inquiry. Bringing UT into the Field. In: *Electronic Journal of the KäTu Symposium on Translation and Interpreting Studies*, MikaEL vol. 12, 29–43. <https://www.sktl.fi/@Bin/65a964ae25090271803bcb6cd7566234/1715336155/application/pdf/1924519/MikaEL12%20-%20Suokas.pdf>

Way, Andy. (2013). “Traditional and Emerging Use-Cases for Machine Translation”, in London, Aslib, *Proceedings of Translating and the Computer* 35. Emerging use-cases for machine translation - ACL Anthology

Zapata, Julián, Polikar, Sébastien. (2023). "La postédition : synonyme de révision?", *Circuit, le Magazine d'informations des langagiers*, Montréal: OTTIAQ. <https://www.circuitmagazine.org/dossier-157/la-postedition-synonyme-de-revision>

Zetzsche, Jost. (2022). "Grades", *The Tool Box Journal* (Issue 22-9-340).

Authors

Tiina Tuominen, Silvia Bacco, Alejandra Jorge, Pascale Elbaz, Isabelle Poff-Pencole.

February, 2024