

Quality Perceptions and Professional Status: The Impact of Extrinsic Information on Translation Editing

Tabea de Wille¹

University of Limerick

Department of Computer Science and Information Systems, Room No: CS2-022 · University of Limerick · Ireland

Montserrat Bermúdez Bausela

Universidad Nacional de Educación a Distancia (UNED)

Facultad de Humanidades de la UNED · Dpto. Filologías Extranjeras y sus Lingüísticas · Despacho 601 · Paseo Senda del Rey 7 · 28040 Madrid · Spain

ABSTRACT

While there is a large body of knowledge on quality in translation and localisation, the question of how quality is *perceived* based on extrinsic factors has so far not been widely investigated. This paper provides some views on translation quality from different theoretical perspectives in the field of Translation Studies and focuses on how extrinsic information on the translators' professional status influences the edits introduced by reviewers to translations in proofreading.

This paper reports primarily on the second part of a two-stage study. In the first stage we found that participants were influenced by extrinsic information when asked to select their preferred translation. This correlation was stronger in participants with less expertise.

In the second stage of the study, presented in this paper, we additionally asked participants to proofread and edit the translations they had selected. We then categorised the changes made (grammar, spelling, meaning, etc.) and conducted frequency analysis and cross-tabulation. Factors for cross-tabulations included for example the participants' level of expertise, and the extrinsic information on the translator's professional status.

Keywords: quality, perception, translation, professional status, localisation, extrinsic factors.

RESUMEN

Si bien es cierto que el conocimiento existente en torno a la calidad en traducción y localización es muy amplio, la cuestión de cómo se *percibe* esta en relación con factores de tipo extrínseco no lo es tanto. Este artículo presenta algunas posturas existentes en torno a la calidad de la traducción según distintas perspectivas teóricas en el ámbito de los Estudios de Traducción y se centra en observar cómo la información de tipo extrínseco que se proporciona al revisor en torno al estatus profesional del traductor le puede llegar a influir a la hora de llevar a cabo el proceso de corrección de una traducción.

En esta línea, el presente artículo detalla los resultados de la segunda parte de un estudio dividido en dos etapas. La primera permitió mostrar cómo los participantes revisores se vieron influidos por factores extrínsecos cuando se les pidió seleccionar la traducción que preferían. Se observó, además, que los factores extrínsecos influyeron en mayor medida en aquellos participantes con menos experiencia. En la segunda parte, detallada a lo largo de este artículo, se solicitó a los participantes que corrigieran e introdujeran modificaciones en las traducciones seleccionadas. Posteriormente, se categorizaron los cambios realizados (en relación con la gramática, la ortografía, el sentido, etc.) a partir de análisis de frecuencia y tablas cruzadas. Algunos de los factores para llevar a cabo la tabulación cruzada incluyeron los niveles de experiencia de los participantes y la información extrínseca sobre el estatus profesional del traductor.

Palabras clave: calidad, percepción, traducción, estatus profesional, localización, factores extrínsecos.

¹ Corresponding author · Email: tabea.dewille@ul.ie

1. Introduction

In the translation seller-buyer relationship, a central issue is not only how quality should be defined, but also the implications of different reviewers with differing levels of measurement knowledge (the individual's ability to assess attribute performance with minimal bias or variance relative to more objective measures (Golder *et al.*, 2012)) evaluating translations. In this paper we reflect on how translation quality (absolute quality) is defined and understood in the light of different quality assessment models and on how to apply them in the evaluation process based on different criteria. We then describe different types of quality attributes and the level of measurement knowledge associated with them.

In order to gain insights into the relationship between reviewer expertise and changes made during translation reviews, we conducted a study in which we asked a group of participant reviewers with different levels of measurement knowledge to choose their preferred translation out of three versions, each one performed by a translator with a different training and expertise in translation, and after that, they were further asked to introduce the edits that they considered necessary so that the final text met their expectations of a 'good' translation.

2. Literature Review

2.1. Translation Quality Assessment (TQA)

Translation quality has always been of primary concern for Translation Studies. However, there is not, so far, a single objective method to evaluate a translation. Furthermore, academia and industry do not always agree on the criteria used to define quality (Drugan, 2013; Chesterman & Wagner, 2002; Pym 2010). One of the problems that lie at the core of assessment in translation is lack of objectivity and value judgements.

Qualitative quality views are strongly linked with different Translation Studies approaches. The concept of equivalence implies the need to talk about the quality of a translation (House, 2015, p. 5). We can distinguish between two different types of equivalence: a more formal linguistic equivalence, which prioritizes both form and content of the source text (ST) (Vinay & Darbelnet, 1958; Catford, 1965; Newmark, 1991); and a pragmatic equivalence, whose focus of attention is no longer the ST, but the recipient's response of the translation. Nida (1964, p. 166) refers to it as 'dynamic equivalence', which may be described as "the closest natural equivalent to the source-language message".

Pragmatic approaches are descriptive in nature, i.e., they focus on describing the process of translation and studying the norms that underlie the translation process. When we understand the norms that prevail in the target system, we also understand the nature of the type of translation being conducted (Toury, 1978; Chesterman & Wagner, 2002; Lambert & Van Gorp, 1985; Hermans, 1985; Lefevere, 1992).

These approaches suggest models that are target-oriented and that focus on the function that the translation is to achieve in its new context of reception. This is the focus of *Skopos* theory, for which the intended function of the target text (TT) will determine the strategy of translation to follow and the quality of the final product (Reiss & Vermeer, 2013). For functionalist Nord (2018, p. 67), evaluating a 'good' translation as being either relatively 'functional' or 'adequate to the purpose' depends on the communicative function it is supposed to achieve. She proposes a translation-oriented textual analysis considering both extratextual (situational) and intratextual (linguistic) factors that help to characterize the function and decide on the most appropriate translation approach.

On her part, functionalist Reiss (2000, p. 3) insists that the comparison between the ST and the TT is essential in order to achieve a balanced judgement. Literary (text types), language (linguistic elements) and pragmatic categories (non-linguistic determinants) are the most reliable way to assure an objective criticism of translations.

For textual and discourse approaches, all texts are evidence of a communicative transaction in a particular social framework. That is why choosing a pragmatic equivalence or a referential one depends on socio-linguistic factors such as the sender's purpose, the function or the recipient of the text (Hatim & Mason, 1990; Neubert & Shreve, 1992). One of the best-known TQA models is the one developed by House (2015), originally coined in 1977 and subsequently revisited in 1997 and 2015. For her, the analysis of variables such as field, tenor, mode and genre, as well as the comparison of textual profiles, provides a useful method to assess the quality of a translation. She points at some constraints dependent on: grammar, lexis, semantics, terminology, phraseology, genre and register, as well as criteria dependent on situational factors. This analysis will help the translator to find out the 'statement of function' of both the ST and the TT that will conclude with the 'statement of quality' of the text. Her model, inestimable as it is, is sometimes a bit more complex than the previous ones to implement, apart from the fact that a 'good' translation for her would be a text that has kept its function equivalent to that of the original and we believe, maybe more in accordance with functionalist theories, that the function of the TT could vary if the translation brief and the recipients' needs require so. In any case, her model is helpful towards orienting the translator's work thanks to its exhaustive analysis.

Also from a textual perspective, it is worth mentioning Williams' (2009) argumentation-centred model. According to William's theory, TQA systems are mainly based on a microtextual level, not macrotextual, which means that they focus on errors at the sub-sentence level, but they do not evaluate the message of the text. His model checks that a translation renders the main argument structure of the original. The analysis and comparison of ST and TT argument structures is an efficient tool to determine the quality of a translation. As part of this model, he considers some of the problems when designing and applying a TQA model, such as levels of seriousness of error (e.g., major, minor, weak point), or multiple levels of assessment (accuracy, target language quality, format, register, situationality, among others). Also interesting is Al-Qinai's (2000) empirical model for TQA, based on several objective parameters. His focus remains on the degree of efficiency of the translation with regard to the syntactic, semantic and pragmatic function of the ST taking into account the cultural frame of both ST and TT.

As we have observed, from a theoretical perspective, quality seems to depend on a wide number of factors, both intralinguistic and extralinguistic.

2.2. Translation Errors

The concept of 'mistake' or 'error' is very broad in Translation Studies as it is linked with the different approaches to translation (Hansen, 2010). Like that, from a functionalist perspective, a translation error is defined according to the purpose of the translation process or product (Nord, 2018, p. 68). From Nord's point of view, being aware of the translation's goal is vital for evaluating the function it needs to achieve in its new context. It is the translator's task to compare the translation brief with the ST analysis in order to foresee the problems that will be likely encountered. Nord goes on to provide a functional classification of translation errors into four categories: pragmatic, cultural, linguistic, and specific. Her model is comprehensive and straightforward. She

provides examples and solutions for each type of problem, classified in a top-down hierarchy, constituting a fundamental source for translators and translation-trainees.

From a textual prism, authors such as Baker (1992) have approached the notion of error from the perspective of lack of equivalence that takes place at different levels: word level, above word level, grammatical level, textual level and pragmatic level. Her study has a special pedagogical importance as she proposes detailed solutions to the various problems that arise due to lack of equivalence in translation on the different suggested levels. Within a textual framework too, we owe another valuable proposal on errors to House (2015, p. 33), who compares the original and the translation based on a number of lexical, syntactic and textual mismatches in relation to the four variables that she analyses. For her, a ‘mismatch’ is an error, which she classifies into ‘covertly erroneous errors’ (lack of functional equivalence) and ‘overtly erroneous errors’ (related to denotative meanings, as well as omissions, additions, substitutions and cases of ungrammaticality and of dubious acceptability). Her model, based on the aforementioned mismatches, is probably more scholarly oriented than professional, since there is a broad number of factors involved that would make it less realistic for the translator to implement in their daily job. However, the fact that the ‘statement of quality’ is derived after a detailed account of mismatches, makes the model a complete one and invites anyone interested in translation to reflect on the relationship between the concept of error and the status of quality.

It is interesting comparing the previous authors’ suggestions with the ideas presented by Williams (2000, pp. 13–14), who defines ‘critical defects’ as those defects impairing translation of the argument macrostructure (i.e., making the translation impossible), while ‘major defects’ will not render the translation unusable, and ‘minor’ would be simply other transfer defects. His view on what constitutes an error is strongly linked to his argumentation-centred TQA model. We believe that weighing how crucial an error is in relation with the ‘usability of the translation’ provides the translator with an unconventional model that not only takes into account microstructure, but also macrostructure and, with it, the core reasoning, grounds of the argument and claims of the text. His classification, invaluable but somehow difficult, contrasts with the simplicity of Brunette’s (2000, p. 179), who cares for a specific terminology regarding TQA and analyses the impact that language errors have on communication, such as mistakes made regarding grammar, spelling, punctuation, and terminology. We find her terminology and explanations very useful for anyone involved in translation.

For the purposes of this paper, and simplifying it considerably, we could say that a mistake is introduced by a participant when there is a lack of equivalence between the ST and TT pragmatically speaking, i.e., when the participant introduces a mistake that changes the overall intention or purpose of the ST, or when the mistake introduced is unacceptable from a professional perspective regarding a more microstructure level, for example, grammar or terminology.

In harmony with the presented literature and in view of the results of our study we have classified the changes introduced by participants according to: grammar, spelling, typography, addition, mistranslation corrected, mistranslation included, omission, untranslated, terminology, style, order, explicitation, naturalness and other (preferential) changes (cf. 4.2). For us, a ‘preferential’ change is a type of edit that does not involve the correction of a translation error, per se, which would prevent the comprehension of the TT, but a type of edit in which the participant chooses a favourite variant.

2.3. Quality Evaluation and Measurement Knowledge

While Translation Studies have for a long time discussed and described translation quality and how to evaluate it, translation buyers are faced with having to make purchasing decisions and obtain quantitative quality measurements without being able to go to these depths of evaluation. The typical buyer/seller relationship in the translation outsourcing process inherently involves a larger or smaller degree of information asymmetry (Akerlof, 1970) where the buyer has lower measurement knowledge than the seller. Measurement knowledge refers to the individual's ability to assess performance of intrinsic attributes with minimal bias or variance relative to more objective measures (Golder *et al.*, 2012). Intrinsic attributes are product-specific and cannot be changed without changing the nature of the product itself (Zeithaml, 1988, pp. 6–7) and include for example faithful conveying of meaning, correct grammar and adherence to customer terminology.

Often clients may have low measurement knowledge due to a lack of familiarity with target language, culture, legal and technical requirements and in many cases this lack of expertise is a major factor in the decision to outsource in the first place (Dunne, 2006, p. 101). As a result, they therefore have to rely on extrinsic cues like price, brand name and level of advertising (Zeithaml, 1988, p. 6) to be able to decide whether the translation meets their business needs and make purchasing decisions and lower levels of measurement knowledge are associated with higher dependence on extrinsic cues for quality evaluation (Golder *et al.*, 2012).

Review and evaluation conducted by humans play a central role in assisting clients in making purchasing decisions. However, this practice is fraught with difficulties, including overzealous reviews consuming buyers' time (Kelly & DePalma, 2009, p. 9; Drugan, 2013, p. 42), informal quality reviews on the buyer's side leading to post-delivery change requests through delayed feedback and overzealous reviewers on the client side, providing undeserved feedback on the quality of the translations (Txabbarriaga, 2009, pp. 1–2).

A key aspect in the review process is how well-defined criteria for review are, as well as the measurement knowledge of the reviewer. While the industry has defined a number of processes and requirements (Kelly & DePalma, 2008, p. 1; Jiménez Crespo, 2011, p. 134; O'Brien, 2012, pp. 59–57), these, similarly to the criteria previously outlined for Translation Studies, still vary quite significantly and are not universally applicable. While they generally show a high level of consensus in some categories like 'language', which commonly include grammar, syntax, punctuation and spelling, as well as, with less consensus, 'typos', fluency and cultural references, agreement is lower for other areas. Categories like 'terminology' and 'accuracy' are generally in use, but with less agreement on the definition of those categories, and 'style' receiving the least consensus (O'Brien, 2012, pp. 60-61).

This discrepancy is not surprising when considering the different types of intrinsic attributes, described by Golder *et al.* (2012). These can be classified according to how easily they can be measured (ambiguous and unambiguous) as well as to customer preferences (homogeneous and heterogeneous). These two variables result in Preference, Universal or Idiosyncratic attributes (Golder *et al.*, 2012, p. 8).

Universal attributes are unambiguous in measurement and homogenous in preference. They are commonly highly codified and can be detected through automation or with low levels of measurement knowledge. For example, spelling errors in a target text could potentially be detected by a child after having completed primary school and having learned the codified spelling rules of the language. They can also be detected with automatic spell checkers to a large extent (García, 2014, p. 432).

Preference attributes are unambiguous in measurement but heterogenous in preference. Examples include project specific requirements like compliance with client terminology. These attributes can also be measured using automated processes like terminology verification software and lend themselves especially well to being defined in client requirement documentation.

Idiosyncratic attributes are ambiguous in measurement and heterogenous in preference. Examples include discourse elements (rhetoric, style, register, ambiguity, implication, allusion, metaphor) or textual properties (coherence, cohesion). These types of attributes commonly pose issues in the review process as they do not lend themselves to being codified into precise rules (García, 2014, p. 432). Resulting from this lack of codification, idiosyncratic attributes are difficult to teach to novices and are instead primarily experience based, tacit knowledge (Dreyfus & Dreyfus, 1980, pp. 7–12). Without the individual linguist being able to make this experience-based tacit knowledge explicit (Eraut, 2000), these idiosyncratic attributes would appear to be what Fawcett describes as quality assessment proceeding "according to the lordly, but completely unexplained whimsy of 'It doesn't sound right'" (Drugan, 2013, p. 37) and are as a result a potential source of disagreement between different linguists.

Overall, objective, systematic and professional quality evaluation requires a certain amount of expertise and knowledge (Havumetsä, 2012, p. 56). This involves primarily domain, linguistic and production knowledge. Students in translation schools might rarely be exposed for long periods to the vocabulary and textual conventions of a domain (Shreve, 2002, p. 163). However, by virtue of having studied a particular language combination, their linguistic knowledge can generally be expected to be high. In addition, they could be expected to have acquired what Muñoz Martín (2009) describes as a part of domain knowledge, and which will be referred to here as production knowledge for easier distinction. Production knowledge for the purpose of this discussion includes translation-specific declarative knowledge such as, for example, conversion of units and measures, using the tools of the trade, converting units and measures. Commonly, although not exclusively, such skills are acquired through formal training programmes and deliberate practice (Pérez-González & Susam-Saraeva, 2012, p. 151). However, Shreve (2002, p. 163) argues that translators "certainly do not have the basis for developing domain-specific production rules, strong methods, for accomplishing the translation task" through translator training programmes. Whether such skills are acquired primarily through training or through working experience, at least linguists who are formally educated in addition to being experienced are likely to have high levels of production knowledge.

Based on the theoretical background in this and the previous chapter, we have designed our study with a focus on the following considerations: While generally, translation studies look at quality from a qualitative perspective, industry buyers additionally require quantitative metrics in order to fulfil their business needs. A challenge here is applying qualitative views of quality to quantitative metrics, in particular for idiosyncratic attributes, as they do not lend themselves to easy definitions to the extent that universal and preference attributes do. By using a category for preferential changes in our study, these idiosyncratic attributes are partially addressed, as far as they are not covered by other categories like style.

With our study design, we aim to mimic the two tasks commonly undertaken by industry reviewers. One, corrections of text with a qualitative focus, which is a task undertaken by our participants. Two, classifying and counting errors, which we have done as part of our analysis.

In order to evaluate the impact of training or training and experience, we have selected three participant groups. Those with no training or experience (novice translators), those with translation training but no experience (advanced translation students) and those with training and experience (lecturers).

These considerations resulted in our hypotheses described in the following section.

3. Methodology

In this paper, our aim was to analyse patterns in quantity and types of edits against the variables:

- The information provided on the professional status of the translator;
- The professional status of the study participants, who acted as reviewers.

By observing the types of edits and how consistently they were made by the majority of participants in a respective group (cf. 3.1.3 to see the demographics of the study in relation to how we distributed the participants in three different groups according to their level of measurement knowledge), we were also able to make observations on absolute quality, in addition to the perceived quality which was the main focus of this study.

Our hypotheses were:

H1: The less experienced or trained the participant reviewer, the more errors they would introduce.

H2: The more experienced or trained the participant reviewer, the fewer preferential changes they would make.

H3: Participant reviewers with higher levels of expertise would introduce a higher number of minor changes regarding style, spelling and punctuation. We believed that if participant reviewers had received sufficient training or were already experts, they would be able to go into deeper detail in their proofreading process and introduce changes related to style, spelling and punctuation, which novice participants might not consider as much.

H4: We also expected major changes made by all groups.

H5: The information provided on the professional status of the translator would influence the amount and the target of the changes. We believed that when presented with more trained or expert translations, participants would introduce fewer changes. On the contrary, when presented with more novice translations, participants would include a higher number of changes. However, we also believed that this information would have less impact in participants with more expertise.

3.1. Description of the study

3.1.1. Text and translation of the text

The text used for the study had been created by one of the authors of this paper and proofread by a native speaker who was trained and experienced in editing texts. Its features were:

- No previous translation was available.

- 13 short paragraphs (two to three sentences each) with individual headings (between two and 13 words each) to avoid as far as possible issues with missing references, context or consistency that may arise from segmenting longer sections.
- General subject matter without requiring high levels of specialisation.

The text was then independently translated by six Spanish translation students and lecturers (two from each group). The groups were:

- Novice translators: first-year translation students who had not yet received training in translation.
- Trained translators: fourth-year translation students who had received in-depth training in translation but had not gained real-life translation experience.
- Expert translators: lecturers at the University Alfonso X el Sabio in Madrid. Both had many (25 and 15 respectively) years of teaching experience in the translation degree and were trained and experienced in judging language quality.

All translators were told that their translations would be used for a study but not what the exact nature of the study was. The translations were evaluated (but not edited) by the researchers to ensure a minimum quality level. In order to qualify, translations had to be mostly complete and not include additional text like translator comments. However, none of the translations were excluded for not meeting this minimum level of quality.

3.1.2. Study design

We then asked different groups of participants, with different levels of measurement knowledge (cf. 3.1.3), to select their preferred version out of three translations (one from each translator group) offered alongside the English ST and edit it so that the final result would meet their expectations of a good translation.

They were presented with 12 scenarios of varying lengths. In four scenarios, they were given no information (NoInfo) on the translator's background (whether it was a novice, trained or expert translator); in four other scenarios, they received correct information about this (GoodInfo); and in the other four, the participants were given incorrect information (BadInfo) on this status.

For students, the data collection was done in a controlled-environment context in the translation laboratory, with the researcher in the same room. Lecturers, on the other hand, completed the reviews in their own time. The reason why methodologies varied across participants is that we felt that students would need supervision and would probably have more questions during the data collection. On the other hand, we were quite certain that lecturers could complete the reviews without supervision. Besides, it would have been difficult to gather all lecturers in the same room to undertake the reviews due to different duties and schedules.

This study was accompanied by a questionnaire in which participants were asked for their opinion on the connection between the professional status of a translator and quality; they were asked to describe what in their own words defines a localisation or translation professional, and to provide a definition of quality.

The part presented in this paper includes the demographics section and questions on the connection between quality definitions and the changes introduced, as well as the connection with professional status.

3.1.3. Demographics

Participants were grouped into three categories, based on their level of training received at that point and their working experience (Table 1).

| Participant grouping | Numbers | Percentage |
|--|----------------|-------------------|
| 1 st /2 nd year students (Novice reviewers) | 11 | 23% |
| 3 rd /4 th year students (Trained reviewers) | 21 | 43.80% |
| Lecturers (Expert reviewers) | 16 | 33.30% |

Table 1. Grouping of participants at UAX.

From the data, we obtained that 43 of 48 respondents self-identified as native speakers of Spanish; the remaining five were also included in the results as they were proficient enough to have been enrolled in the translation course.

None of the 1st and 2nd year students had professional translation experience, and the category 'less than 12 months of experience' was selected primarily by 3rd and 4th year students, while those who had selected 5 to 30 years or up to 5 years of experience were predominantly part of the lecturer respondent group. Some of the lecturers also had extensive training in translation, specialist knowledge in linguistics and 79% of the lecturers had received training in conducting reviews (academic or non-academic).

4. Results

4.1. Quality Definitions

In order to determine a potential link between the participants' definition of quality and changes made to translations during review, we asked participants to provide, in their own words, a definition of quality. Of the total 48 responses, 13 were excluded as they did not provide a definition of quality. Of the remaining 35 responses, seven mentioned process-related aspects of quality like specifications or requirements (3), involving more than one translator (1), timely delivery (1), basing the translation on the original text (1) and revision (1).

Our third step was to assign labels depending on specific mentions of quality markers, based on the categories used for evaluating the changes made (Table 2). The majority of the definitions included faithfulness to the original as one of the most important criteria for a translation to be qualified as good (Figure 1).

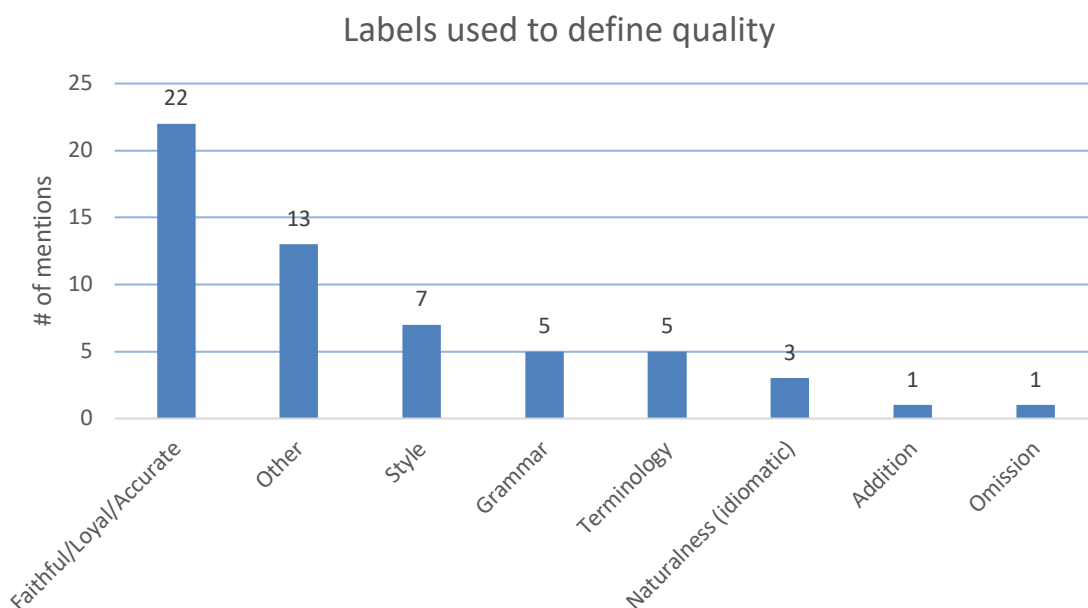


Figure 1. Specific mentions of quality markers.

| Label | Explanation | Example of definition provided by participants |
|-------------------------|---|--|
| Faithful/Loyal/Accurate | Same meaning as the original. | A good translation would be one that is closest in meaning to the original text. |
| Other | Additional remarks: final product quality, culture and context, topic, meeting client specifications, purpose, adaptation, coherence, understanding the text. | A translation is good if it is able to achieve the objectives the original text was seeking. |
| Style | Stylistic issues of great importance for accomplishing a good translation. | Providing the same meaning in the target language but respecting the style of the source language. |
| Grammar | Correlating quality with the lack of grammar mistakes. | To have a text that is free of linguistic errors, that sounds natural and that is accurate based on the ST. |
| Terminology | The use of the appropriate terms. | You have to take into account several things such as terminology, grammar, etc. |
| Naturalness | Focusing on the effect of the TT. | The effect of the translated text must be the same as with the original text. |
| Addition | Quality requires that the translation does not add more information than the one stated in the original. | Convey faithfully (without adding or subtracting) the message of the source language in a natural manner for the target. |
| Omission | A good translation would be one that has translated all the elements of the original, not leaving any elements out. | Convey faithfully (without adding or subtracting) the message of the source language in a natural manner for the target. |

Table 2. Categories used for evaluating the changes made.

4.2. Study results

We analysed the changes introduced by each individual participant for each paragraph. Changes made by participants were labelled as shown in Table 3 based on both the literature referenced in this paper and emerging from the results.

| Label | Definition | Example |
|--------------------------|--|---|
| Grammar | Changes related to the structure of the sentence and classes of words | ST: You have been accepted to... Has sido Te han admitido... |
| Spelling | Changes in which the sequence of letters is wrong and issues that have to do with orthography. | ST: as libraries and other practitioners in the industry. como las bibliotecas bibliotecas u otros trabajadores del sector. |
| Typography | Changes that deal with punctuation | ST: Establish priorities in what to study Establezca prioridades (en lo relativo) a qué estudiar |
| Addition | The participant adds information that is not, as such, in the ST or in the original translation. | ST: and also slowly chips away at the overall pile of tasks. ...y también a ir avanzando poco a poco con la totalidad de las tareas <u>para que no se te acumulen</u> |
| Mistranslation corrected | There was a mistake in the selected translation that affected the overall meaning of the sentence and the participant corrected it. | ST: ...getting into academia. ... entrar en la Escuela Mayor. <u>continuar con tu formación académica.</u> |
| Mistranslation included | The respondent corrected something that was right either in meaning or form in the original translation and changed it into something wrong. | ST: Show up Mantente en contacto <u>Ponerse al descubierto</u> |
| Omission | The participant omitted to translate something, i.e., they left some part of the sentence out. | ST: ...what you want to get out of it qué quiere obtener de ello |
| Untranslated | The participant 'reverts' to a term in its source language. | ST: Succeeding in the MSc Cómo tener éxito en el máster <u>MSc.</u> |
| Terminology | Changes related to the selection of specific vocabulary. | ST: ...getting credentials in order to get promoted into a specific role... |

| | | |
|---------------------|--|--|
| | | ...conseguir una acreditación necesaria <u>referencias</u> para lograr un ascenso a un puesto concreto... |
| Style | Changes related to stylistic issues and register (lowering it or making it more formal than the original translation) | ST: If you do come across something interesting, don't forget to share! Si encontraste <u>topas</u> <u>con</u> algo interesante, ¡no olvides <u>compartirlo!</u> |
| Order | Changes related to the placement of some elements of the sentence in different places. | ST: the wide diversity of students participating on all levels la amplia diversidad, <u>a todos los niveles</u> , <u>de los</u> <u>estudiantes</u> que participan en todos los niveles : |
| Explicitation | When a pronoun is replaced with its referent, a common word is replaced with a proper name, or the reviewer clarifies an idea repeating an element of the sentence. | ST: and when and how you will be able to do your research and writing. y cómo podrás hacer tu investigación y la redacción <u>de la tesina.</u> |
| Naturalness | Changes made to improve how idiomatic the original translation is. | ST: Show up No faltas <u>dejes de asistir</u> |
| Other (preferences) | Preferential change, usually consisting in replacing one word or expression with another one belonging to the same part of speech. We call it 'preferential' because it does not correct an error in the original translation. It is just a matter of choosing a favourite option over the one provided by the original translation. | ST: The internet is full of possibilities, as are libraries. Internet está lleno de posibilidades, al igual de la misma <u>manera</u> que las bibliotecas |

Table 3. Type of changes made by participants.

The following sections first show the numbers of changes made, then the types of changes and finally, the mistakes introduced. All the results that we provide in this paper are adjusted sums. This means that we have taken the average number of changes introduced by each group of participants and related it to the number of

responses obtained per group as well as the number of times a particular translation was chosen as the best version, if applicable.

4.2.1. Numbers of changes made per participant group

With 81% of the changes, novice reviewers (1st/2nd year students) introduced more changes than any other participant group, even though only 23% of the participants belonged in this category (Table 1). Trained reviewers (3rd/4th year students) introduced 13% of the changes; expert reviewers (lecturers), only 6%.

This trend was especially pronounced in NoInfo and BadInfo scenarios (Figure 2).

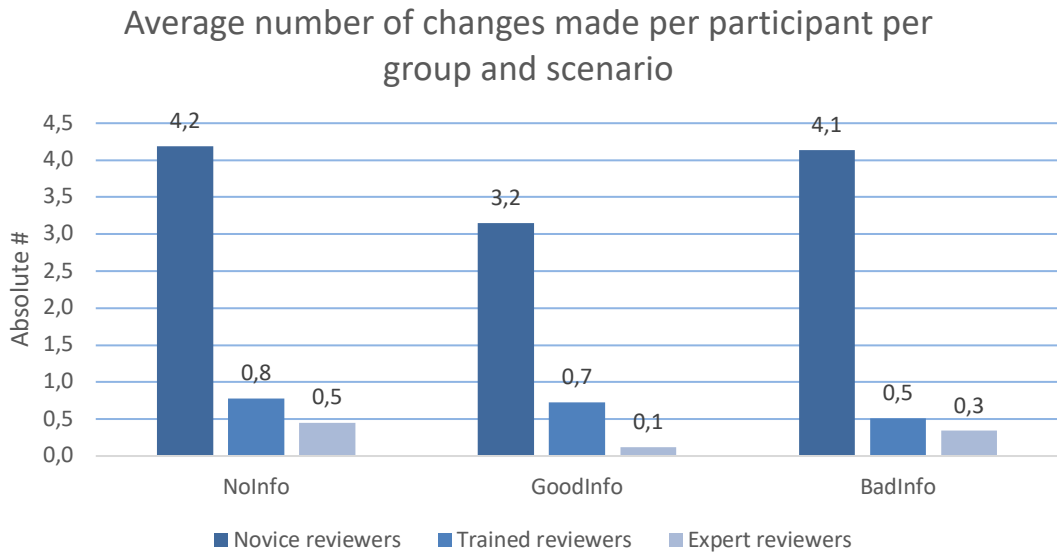


Figure 2. Average number of changes made per participant per group and scenario.

When reviewers were given no information on the translator’s background (NoInfo), translations provided by expert translators (lecturers) were chosen as the best most frequently (44%), but they register the least percentage of changes (29%). Trained translations (4th year students’) were changed most frequently (36%), followed by novice translations (1st year students’) (35%) (Figure 3).

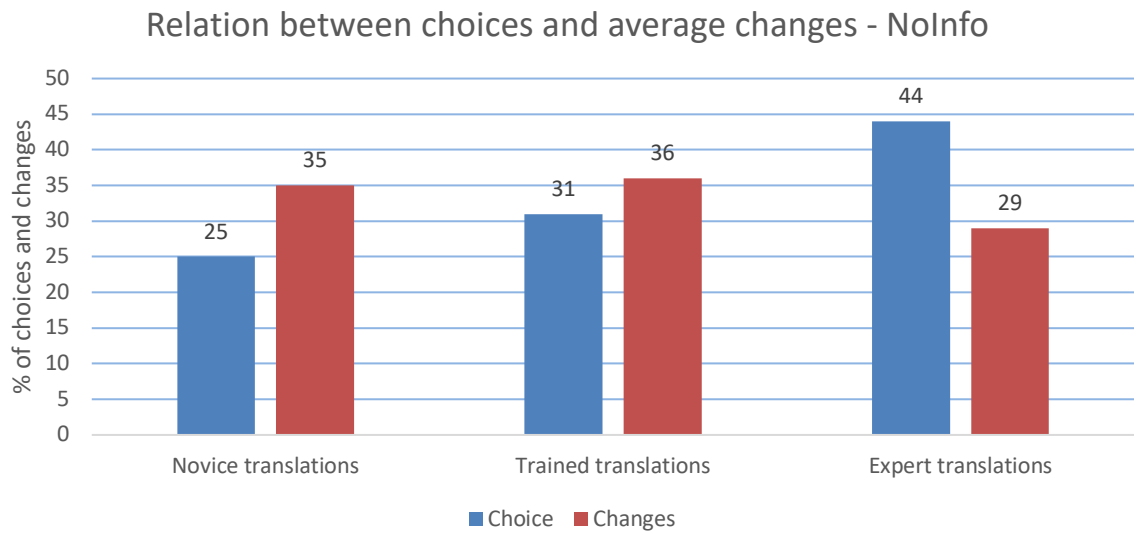


Figure 3. Relation between choices and average changes given NoInfo.

While novice reviewers introduced more changes to all the translations in general, they in particular did so to trained translations. Expert reviewers primarily introduced changes to novice translations (Figure 4).

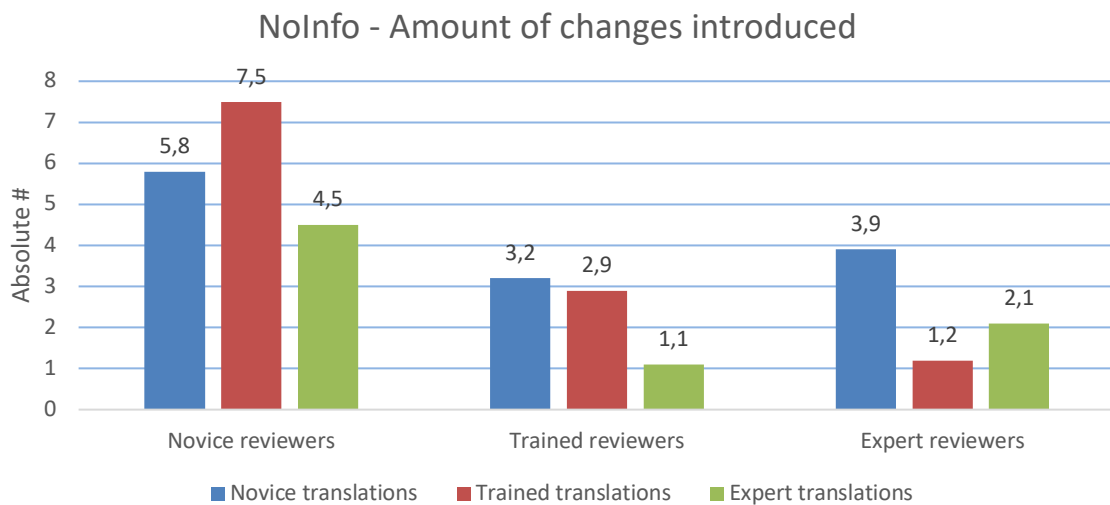


Figure 4. Amount of changes introduced given NoInfo.

When they were given good information on the translator’s background (GoodInfo), all groups selected the expert translators’ versions most frequently (71%) and introduced the highest percentage of changes (58%) to these translations. As mentioned, in NoInfo scenarios, expert translations selection was 44% and they registered the lowest number of changes (29%). So, we can say that the information provided to participants does make a difference both in their choices and the target or focus of their changes. Let us remember that all the results are adjusted.

It should be noted that novice translations received 27% of the changes (second position), even though they were picked as the best translation least frequently (13%) (Figure 5).

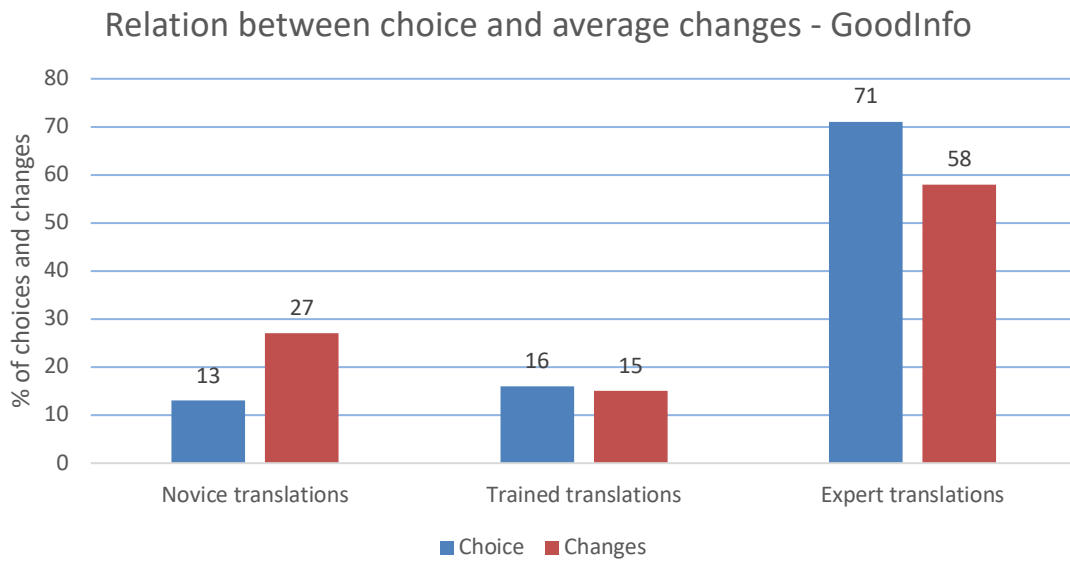


Figure 5. Relation between choices and average changes given GoodInfo.

When correct information was provided on the translators' background (GoodInfo), again novice reviewers introduced more changes to the translations, while it is noticeable how expert reviewers hardly introduced any changes. Novice and trained reviewers introduced more changes to novice translations. Experts had not picked the novice versions as the best at all. As a result, their changes were focused on the expert translators' versions (Figure 6).

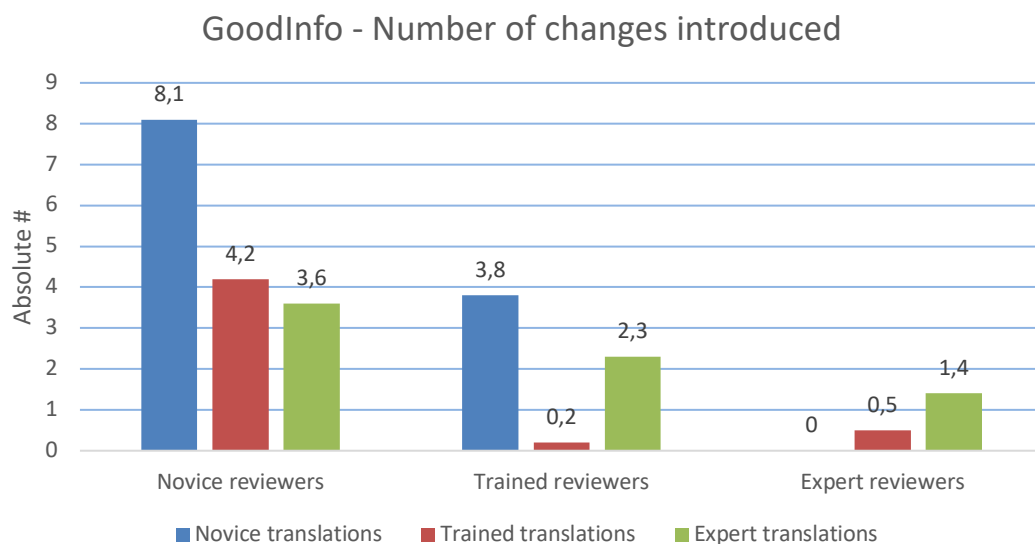


Figure 6. Amount of changes introduced given GoodInfo.

When participants were given wrong information on the translator's background (BadInfo), the differences in selection and the changes made were in general less pronounced, indicating overall uncertainty due to the discrepancy between intrinsic attributes and extrinsic information (Figure 7).

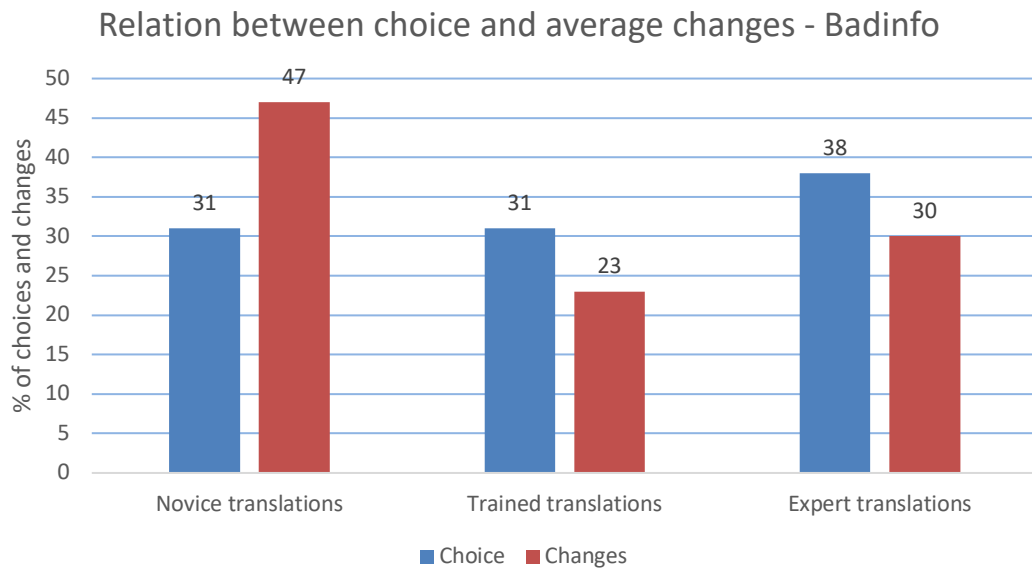


Figure 7. Relation between choices and average changes given BadInfo.

The expert translation (Truth) was picked most frequently (38%), but markedly less so than in GoodInfo scenarios (71%).

However, the expert translators’ version (Truth) only received 30% of the changes. Most frequently, novice translations (Truth) (47%) were the focus of the changes.

In BadInfo scenarios, again novice reviewers introduced more changes (Figure 8).

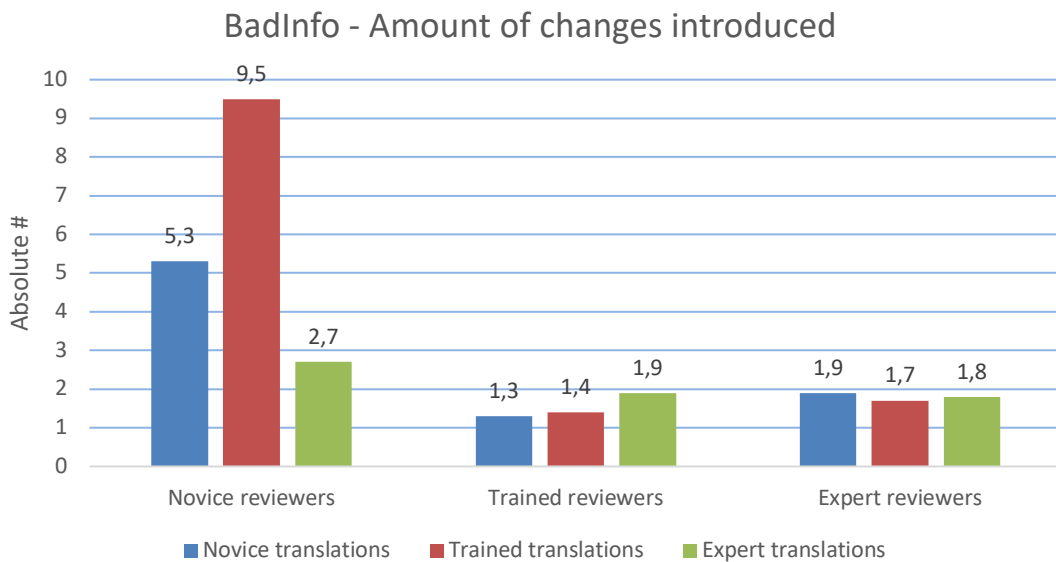


Figure 8. Amount of changes introduced given BadInfo.

4.2.2. Types of changes made across all participant groups

Grammar and terminology changes were most frequently introduced by all participants, followed by style and addition. Among the changes with overall lower numbers, changes aimed at providing a more idiomatic translation, with a more adequate style in the target language were quite high.

Interestingly, we could observe two opposite tendencies: changes introduced that omitted part of the text and changes that added more information to it (Figure 9).

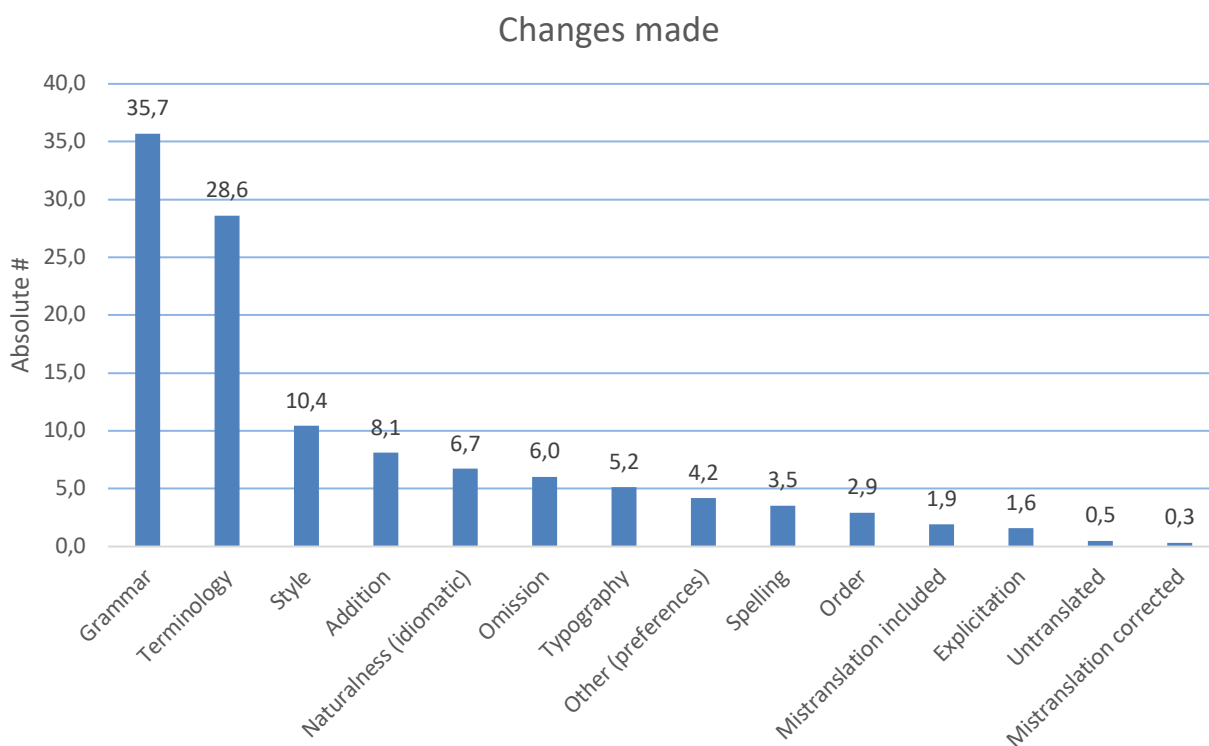


Figure 9. Number of changes made across participant groups.

Regarding the information provided (three different scenarios) and the type of translations being edited, we have established the following patterns concerning types of changes (Table 4). We give next to them absolute numbers; we only include the first five most common types of changes and we do not show changes with absolute numbers below 0.5:

| | Novice translations | Absolute # | Trained translations | Absolute # | Expert translations | Absolute # |
|----------|----------------------------|-------------------|-----------------------------|-------------------|----------------------------|-------------------|
| NoInfo | grammar | 4.4 | grammar | 5.1 | grammar | 4.2 |
| | terminology | 3.7 | terminology | 4.3 | terminology | 2.4 |
| | style | 1.6 | style | 1.6 | typography | 1.0 |
| | addition | 2.0 | typography | 1.0 | addition | 1.1 |
| | typography | 1.1 | addition | 0.8 | omission | 0.8 |
| GoodInfo | grammar | 2.8 | grammar | 2.4 | grammar | 6.5 |
| | terminology | 2.6 | typography | 0.5 | terminology | 4.7 |
| | style | 0.8 | -- | -- | omission | 1.6 |
| | omission | 0.6 | -- | -- | naturalness | 1.5 |
| BadInfo | terminology | 6.0 | grammar | 3.6 | grammar | 4.1 |
| | grammar | 5.4 | terminology | 2.3 | terminology | 2.3 |
| | style | 2.2 | addition | 1.1 | spelling | 1.1 |

| | | | | | |
|----------|-----|-------------|-----|-------------|-----|
| spelling | 1.1 | naturalness | 1.1 | style | 1.1 |
| omission | 1.1 | style | 1.1 | naturalness | 1.0 |

Table 4. Patterns in the type of changes introduced for the three scenarios.

4.2.3. Type of changes introduced by participant groups

This could also be studied with a different focus: in relation to the type of changes that each participant group tended to introduce the most (Table 5). The results were quite equally independently of the information provided.

| Reviewer participants | Type of changes | Absolute # |
|-----------------------|-----------------|------------|
| Novice reviewers | Grammar | 24.9 |
| | Terminology | 18.2 |
| | Style | 6.5 |
| | Naturalness | 4.7 |
| | Addition | 4.6 |
| Trained reviewers | Grammar | 7.5 |
| | Terminology | 4.8 |
| | Style | 2.2 |
| | Addition | 2.0 |
| | Omission | 1.5 |
| Expert reviewers | Grammar | 6.2 |
| | Terminology | 5.7 |
| | Omission | 1.9 |
| | Style | 1.9 |
| | Addition | 1.7 |

Table 5. Type of changes per participant group.

4.2.4. Mistakes introduced in the changes and its relation to the participant background

As we can observe, novice reviewers introduced a larger number of mistakes than the other two groups (Figure 10).

Average amount of mistakes introduced

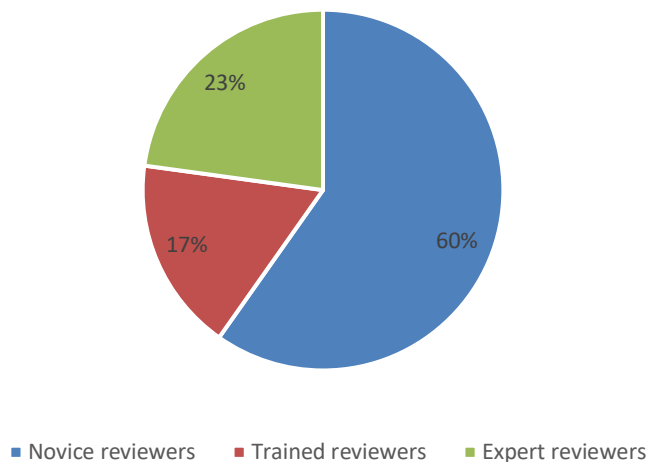


Figure 10. Average amount of mistakes introduced across all groups.

The information provided to the participants on the translators' background made no difference as to the number of mistakes introduced in the translations.

We could observe a total of 19 mistakes introduced by all participant groups:

- Seven of them were problems related with meaning.
- 10 of them had to do with spelling. Out of them, three concerned orthography and seven concerned misplaced letters in the word.
- Two of them dealt with grammar problems.

Also, most of the mistakes introduced by novice reviewers were related to meaning. In second and third positions were spelling and grammar mistakes respectively. We distinguished between spelling problems that we called 'probably involuntary', since they usually involved a misplaced letter; and spelling problems that dealt with more serious issues such as orthography. In this sense, novice reviewers introduced both involuntary spelling mistakes and orthographic ones.

Mistakes introduced by trained reviewers belonged in two of the cases to wrong meaning, one probably involuntary spelling mistake and one grammar mistake.

In the expert reviewers' group, one of the mistakes concerned meaning, three were probably involuntary spelling mistakes, and one was a spelling mistake (orthography).

5. Discussion

5.1. Number of changes introduced overall

Novice reviewers introduced more changes in all situations, independently of the information provided on the professional status of the translator. After them, and at a great distance, came the changes introduced by trained reviewers, while the last of all were the experts, with the lowest percentage of changes being introduced.

5.2. Nature of changes introduced

Types of changes introduced were quite similar across the three participant groups and independently of the three possible scenarios. We could perhaps observe more variety in the changes introduced by experts.

All participants tended to introduce changes related to grammar and terminology, after which, and at a great distance, we identified other types of changes, in particular:

- Novice reviewers: style, naturalness and addition.
- Trained reviewers: style, addition and omission.
- Expert reviewers: omission, style and addition.

5.3. Professional status of the translator and changes introduced

The results also show that the information provided on the professional status of the translator had an impact, not so much on how much was edited, but on the focus of the type of translation being corrected. Given no information, the target of the changes were trained translations; when good information was provided, the target were the expert ones; and, finally, when wrong information was provided, the focus of the changes were novice translations.

| | Choice | Changes |
|-----------------|---------------|----------------|
| NoInfo | Expert (44%) | Trained (36%) |
| GoodInfo | Expert (71%) | Expert (58%) |
| BadInfo (Truth) | Expert (38%) | Novice (47%) |

Table 6. Relationship between choices and changes in the three scenarios.

Overall, novice translations were modified the most (36%), followed by expert translations (34%) and trained translations (30%).

5.4. Mistakes introduced

As we have observed, we found that the less experienced or trained the participant, the more errors they introduced in their own changes. Also, the type of mistakes they introduced was relevant for this study, as we could see that novice participants tended to include errors related mainly to meaning and orthography, while experts, on the other hand, introduced mistakes that were very likely involuntary spelling errors as they mainly consisted in misplaced letters.

5.5. Quality definitions and changes introduced

For the reviewers, in general, the most important criterion that the translation needs to meet in order to be qualified as good would be a text that renders the same meaning as the original, that it is faithful and loyal to the original. In this sense, 22 definitions of quality established faithfulness to the original as the most important factor, and only two definitions mentioned adaptation as criterion for a good translation. We could draw from this that reviewers are probably more inclined to meet a more formal and traditional type of linguistic equivalence to a communicative one.

After that, reviewers estimated that the right terminology and the proper grammar in the translation are paramount for a TT to be qualified as good. Like this, there is a connection between the nature of the changes made and how they defined quality in a translation, since grammar and terminology were the two most common type of changes.

From observing the edits made, we could also conclude that the aim of reviewers introducing grammar and, particularly, terminology changes, was to correct, according to their criteria and opinion, a translation that lacked the appropriate or most adequate equivalent in meaning, form, and style. In this regard, for reviewers, terminology, grammar, typography, spelling and style all add up to create a TT that is faithful to the original.

Stylistic changes and those related to naturalness were also particularly abundant. And this was also the case in the definitions provided by participants. In 10 cases they mentioned style and naturalness as of paramount importance to provide a good translation.

With “Other” we labelled other definitions that included aspects not mentioned or not classified under the labels created for such purpose. Those aspects referred to culture, objectives met, adapting the text, requirements, context, revision, writing skills, resources, intelligibility and effect.

As an anecdotal detail, we could mention that one definition of quality required that the translated text did not add more information than the one included in the original, while another definition stated that no elements of the sentence should be left out. However, both addition and omission have been two common types of changes among participants in the sense that they either added some elements when editing their preferred translation, or deleted them from the original translation. We believe that they did so with the aim to elaborate a more natural final text.

However, and apart from that, we can conclude that there was coherence between the quality definitions provided by participants and the type of changes that afterwards they made to the original translations.

5.6. Impact or influence of the length of the paragraphs

We could observe that the length of the paragraphs did not have an impact regarding the number of changes introduced. Both were equally changed.

6. Conclusion

Let us now revisit our initial hypotheses in order to discuss to what extent they match the results of the study and draw some conclusions:

- H1: The less experienced or trained the participant reviewer, the more errors they would introduce: This hypothesis has been fulfilled since most of the errors introduced belonged to novice participants.
- H2: The more experienced or trained the participant reviewer, the fewer preferential changes they would make: There were, in general, very few preferential changes. These were not outstanding in any group. Preferential changes consisted in choosing one term, verb, expression, prepositional phrase, adverbial phrase, etc. over another one. Order was also a type of preferential change.
- H3: The more experienced or trained the participant reviewer, the more minor changes regarding style, spelling and punctuation they would introduce: Apart from style, which was a type of change quite frequently made by all groups, expert reviewers also introduced more changes related to addition, omission and spelling, which shows more variety in the changes introduced by experts.
- H4: We expected major changes made by all groups: This was particularly true in the case of the novice participant group and at a great distance from the other two groups, while the least number of changes were introduced by expert reviewers.
- H5: The information provided on the professional status of the translators had an impact on how much was edited and the target of the edits. We could observe that it was not a matter of how much was edited, but the target of the edits what really mattered to participants. When NoInfo was provided on this, changes affected trained translations mainly; when GoodInfo was provided, the focus were the expert translations; and, finally, when BadInfo was provided, the target of the changes were novice translations.

From these, we could derive some conclusions. To begin with, we have noticed that there were patterns in which some translations from one source were edited more than others, in particular expert translations, followed by novice translations. Also, that the information provided on the professional status of the translator *was* a determining factor regarding the target of the edits. When novice participant reviewers knew the source, they were more willing to introduce changes to novice translations (maybe knowing that it was a novice translation gave them the freedom and authority to try to improve it), while expert participant reviewers devoted their efforts to the expert translations and introduced only slight changes to improve the translation. On the other hand, when given wrong information, confusion among novice reviewers seemed to be the rule since the focus of their edits was, this time, the trained translations; while for the trained and expert reviewers, the edits introduced were equally distributed among the three types of translations. We can observe how training and professional experience are key factors in order to avoid being biased by extrinsic cues.

Also interesting to note is the fact that novice participant reviewers were more prone to introduce more changes and they did so independently of the information provided on the status of the translation. On the opposite side, trained participant reviewers and, even more so, expert participant reviewers introduced the least number of changes. A plausible conclusion could be that the more trained the reviewer is, the more thoughtful he or she is before introducing edits to the translation. Novice reviewers, on the other hand, may be more eager to introduce as many changes as possible in their ‘debut’ of editing a translation. Therefore, it is not the number of changes introduced that make a translation ‘better’, it is the necessary and pertinent changes being introduced to correct or improve it that really matters. Here, translation training and expertise help to develop this competence according to our findings.

Finally, we can conclude that similar edits were made consistently. This was the case among all groups and in all three scenarios. We mentioned before that this could raise the issue of absolute quality compared to the perceived quality. In this sense, we have been able to observe the majority vote on some changes that needed be made.

About the authors

Dr. Tabea De Wille holds an MA (Magister Artium) in German and English Linguistics from the University of Erlangen-Nürnberg, Germany and an MSc Multilingual Computing and Localisation from the University of Limerick, Ireland. She has completed her PhD studies at the University of Limerick, Ireland within CNGL II, where she has examined perceived quality in the context of crowdsourced localisation.

Tabea has in the past worked in the localisation industry, primarily in video games localisation and is currently a lecturer at the University of Limerick where she teaches in a range of localisation, software internationalisation and translation technology modules as well as modules related to video games and project management. Tabea is the director of the Localisation Research Centre at the University of Limerick.

Montserrat Bermúdez Bausela is a lecturer in English at UNED (National Distance Education University) where she teaches Legal English, English for Specific Purposes (ESP) and Specialised Translation in different degrees as well as in the MA in the European Union. She has also taught Linguistics, Translation, Computer-Assisted Translation Tools and Software Localisation at Universidad Alfonso X el Sabio.

She holds a degree in English (Universidad de Valladolid), a PhD in Philology (UNED), an MSc in Localisation (University of Limerick) and an MA in Specialised Translation (Universidad de Valladolid). Her research interests are English for Specific Purposes, Translation Studies and Corpus Linguistics, among others. She is the author of a number of articles in leading journals on Translation Studies.

She has recently translated from English into Spanish the book by Sari Nusseibeh, *Once upon a Country: a Palestinian Life* (*Érase una vez un país: una vida palestina*), published by Berg Institute.

References

- Akerlof, G. A. (1970). The Market for “Lemons”: Quality Uncertainty and the Market Mechanism. *The Quarterly Journal of Economics*, 84(3), 488–500.
- Al-Qinai, J. (2000). Translation Quality Assessment: Strategies, Parametres and Procedures. *Meta*, 45(3), 497-519.
- Baker, M. (1992). In Other Words. A Coursebook on Translation. London & New York: Routledge.
- Brunette, L. (2000). Towards a Terminology for Translation Quality Assessment. *The Translator*, 6(2), 169-182.
- Catford, J. (1965). A Linguistic Theory of Translation. Oxford: Oxford University Press.
- Chesterman, A., & Wagner, E. (2002). Can Theory Help Translators? A dialogue between the Ivory Tower and the Wordface. Manchester, UK & Northampton, MA: St. Jerome.
- Dreyfus, S.E. & Dreyfus, H.L. (1980). A five-stage model of the mental activities involved in directed skill acquisition. Operations Research Center.
- Drugan, J. (2013). Quality in Professional Translation. Assessment and Improvement. London & New York: Bloomsbury.
- Dunne, K.J. (2006). Putting the cart behind the horse: Rethinking localization quality management. In K.J. Dunne (Ed.), *Perspectives on Localization*. John Benjamins Publishing Company: Amsterdam/Philadelphia, 95–117.
- Eraut, M. (2000). Non-formal learning and tacit knowledge in professional work. *The British Journal of Educational Psychology*, 70, 113–136.
- García, I. (2014). Training Quality Evaluators. *Revista Tradumàtica: tecnologies de la traducció*, 12, 430–436.
- Golder, P. N., Mitra, D., & Moorman, C. (2012). What Is Quality? An Integrative Framework of Processes and States. *Journal of Marketing*, 76, 1–23.
- Hansen, G. (2010). Translation ‘errors’. In Y. Gambier & L. van Doorslaer (Eds.), *Handbook of Translation Studies: Volume 1* (pp. 385–388). Amsterdam: John Benjamins.
- Hatim, B. & Mason, I. (1990). Discourse and the Translator. London: Longman.
- Havumetsä, N. (2012). The Client Factor: A Study of Clients’ Expectations Regarding Non-Literary Translators and the Quality of Non-Literary Translations. Unpublished thesis, University of Helsinki.
- Hermans, T. (1985). Translation Studies and a New Paradigm. In T. Hermans (Ed.), *The Manipulation of Literature. Studies in Literary Translation* (pp. 7–15). London: Croom Helm.
- House, J. (2015). Translation Quality Assessment: Past and Present. London & New York: Routledge.
- Jiménez Crespo, M.Á. (2011). From many one: Novel approaches to translation quality in a social network era. *Linguistica Antverpiensia New Series - Themes in Translation Studies: Translation as a Social Activity - community translation 2.0*, 2, 131–152.
- Kelly, N. & DePalma, D.A. (2008). Non-Standard Views on Industry Standards. Common Sense Advisory: Lowell, Massachusetts.
- Kelly, N. & DePalma, D.A. (2009). Eliminating Roadblocks to Translation Quality: Practical Steps to Improve Translated Output. Common Sense Advisory: Lowell, Massachusetts.

- Lambert, J., & Van Gorp, H. (1985). On Describing Translations. In T. Hermans (Ed.), *The Manipulation of Literature. Studies in Literary Translation* (pp. 42–53). London: Croom Helm.
- Lefevere, A. (1992). *Translation, Rewriting and the Manipulation of Literary Fame*. London & New York: Routledge.
- Muñoz Martín, R. (2009). Expertise and environment in translation. *Mutatis Mutandis*, 2(June), 24–37.
- Neubert, A., & Shreve, G. M. (1992). *Translation as Text*. Kent: Kent State University.
- Newmark, P. (1991). *About Translation*. Clevedon: Multilingual Matters.
- Nida, E. (1964). *Towards a Science of Translation*. Leiden: Brill.
- Nord, C. (2018). *Translating as a Purposeful Activity: Functionalist Approaches Explained*. Manchester: St. Jerome Publishing.
- O'Brien, S. (2012). Towards a Dynamic Quality Evaluation Model for Translation. *The Journal of Specialised Translation*, 17, 55–77.
- Pérez-González, L. & Susam-Saraeva, Ş. (2012). Non-professionals Translating and Interpreting. *The Translator*, 18(2), 149–165.
- Pym, A. (2010). *Exploring Translation Theories*. London & New York: Routledge.
- Reiss, K. (2000). *Translation Criticism: The Potential and Limitations*. (E.F. Rhodes, Trans). Manchester: St Jerome.
- Reiss, K., & Vermeer, H. J. (2013). *Towards a General Theory of Translational Action: Skopos Theory Explained*. (C. Nord, Trans.). London & New York: Routledge.
- Shreve, G.M. (2002). Knowing translation: cognitive and experiential aspects of translation expertise from the perspective of expertise studies. In A. Riccard (Ed.), *Translation Studies: Perspectives on an Emerging Discipline*, University Press: Cambridge, 150–171.
- Txabarriaga, R. (2009). *Best Practices for Client Review Processes*. Common Sense Advisory: Lowell, Massachusetts.
- Toury, G. (1978). The Nature and Role of Norms in Literary Translation. In J. S. Holmes, J. Lambert & R. Van den Broeck (Eds.), *Literature in Translation: New Perspectives in Literary Studies* (pp. 83–100). Leuven: Acco.
- Vinay, J. P., & Darbelnet, J. (1958). *Stylistique Comparée du Français et de L'Anglais. Méthode de traduction*. Paris: Didier.
- Williams, M. (2009). Translation Quality Assessment. *Mutatis Mutandis*, 8(1), 3-23.
- Zeithaml, V. A. (1988). Consumer Perceptions of Price, Quality, and Value: A Means-End Model and Synthesis of Evidence. *The Journal of Marketing*, 52(3), 2–22.