Towards Defining a Specific Text as a Valid Instrument in Translation Quality Evaluation Studies: A Conceptual Paper

Mehrnoush Norouzi
Islamic Azad University
Central Tehran Branch
Iran.

ABSTRACT

This conceptual paper is an attempt to present and highlight the features required for the selected text as an instrument in translation studies which researchers should consider .This is a step towards making the translation studies more scientific. To fulfil the paper`s purpose, the author discussed some researches findings and based on the some findings, such as Conde`s (2009) ones, tried to present a standard framework to researchers to select a text as the instrument for translation quality assessment in quantitative translation studies. Finally the author brought some features up briefly named the length of the text and the readability information of the text resulted from reasons discussed in details in this paper.

1. Introduction

Translation studies encompass vast domains, familiarly many of which deal with translation quality. Regularly in many experimental studies, translation quality is applied as a variable. In this respect research participants are given different texts to translate and consequently the translated texts are assessed by the evaluators to provide required data for analyzing. In this paper the author intends to draw the reader's attention that, since in many studies the participants are required to present an appropriate and assessable translations to the researcher for his/her investigative purposes in the limited time and special research conditions, the translation quality assessment *text* should have specified features and some cases about the assessment *text* should be considered before head.

2. Defining 'Text'

Richards, Platt & Platt (1992,p.378) define text as a piece of spoken or written language. A text may be considered from the point of view of its structure and/or its functions, e.g. warning, instructing, carrying out a transaction. A full understanding of a text is often impossible without reference to the context in which it occurs. A text may consist of just one word, e.g. DANGER on a warning sign , or it may

be of considerable length, e.g. a sermon, a novel, or a debate.

In the other words , Beaugrande and Dressler (1981,p3) explain: A text will be defined as a communicative occurrence which meets seven standards is not considered to have been satisfied, the text will not be communicative. Hence, non-communicative texts are treated as non-texts.

Later Beaugrande and Dressler (1985/9,p.10) presented another definition of text: We can define text, in the simplest way perhaps, by saying that it is language that is functional. By functional, we simply mean language that is doing some job in some context, as opposed to isolated words or sentences that I might put on the black board.(These might also be functional, of course , if I was using them as linguistic examples).

Halliday and Hasan (1976.p.2) say: A text is best regarded as a semantic unit: a unit not of from but of meaning. Thus it is related to a clause or sentence not by size but by realization ,the coding of one symbolic system in another. A text does not consist of sentence ;it is realized by, or encoded in sentences.

Later Halliday and Hasan (1976,p.23) explain: The concept of cohesion can therefore be usefully supplemented by that of register, since the two together effectively define a text. A text is a passage of discourse which is coherent with respect to itself, and therefore cohesive.

Tony Bex declares (2002,p.78) "text is considered to be that set of visual linguistic Symbols which are capable of being employed in the construction of meaning."

In newest point of view (Kelly R.Iverson, 2015) asserts in modern linguistic studies the word text can refer to material both oral and graphic/written. However, I address the written text in this paper as mean of attention, which is used in many experimental translation studies.

3. Translation Quality Assessment

Translation quality assessment is profoundly important in translation studies. In this regard House (2015, p.1) declared "Translation quality assessment can thus be said to be at the heart of any theory of translation."

Evaluating the quality of translation remains one of the most difficult areas in the study of translation. According to Gharacheh (2005) "the main problem seems to reside in how to express quality or what measure should be used for the quality of a translation"(p.20). Some methods have been introduced by translation scholars during the past decades as a way to evaluate the quality of translated text based on approaches and theories .House (2015) divided the approaches as Response-based approaches including Behavioristic views and Functionalistic , "skopos"-related views, text and discourse-oriented approaches including descriptive translation studies ,philosophical and socio-cultural, socio-political approaches; and linguistically oriented approaches . Based on any of approaches, models and rubrics have been presented to assess the quality of translation.

3.1 Translation Quality Models

The first step to assess the quality of translation is providing a model as a framework .There have been some TQA models by theorists such as ,Van den Broeck's (1985) model of TQA focuses on literary translation and aimed to establish the degree of `functional equivalence between source and target texts. D`Hulst's (1996) model aimed to assess texts with

special goals, trying to create the same function on source and target cultures.

Al-Qinai (2000) model that can be empirically tested for analyzing the linguistic and situational peculiarities of source and target texts in the pre-translational phase and post translational assessment of target text quality. And the most famous and prior model specifically named Translation Quality Assessment by House in 1977.

House's Model of TQA

Juliane House presented a model named specifically Translation Quality Assessment in 1977 ,later in 1997 and lately in 2015 revised it. According to House (2006) "Translation can then be defined as the replacement of a text in a source language by a semantically and pragmatically equivalent text in a target language. An adequate translation is thus a pragmatically and semantically equivalent one." In this regard Baker and Saldanha (2009,p.224) state: House proposed a model based on pragmatic theories of language use; this model provides for the analysis of the linguistic-situational   particularities of source and target texts, a comparison of the two texts and the the resultant assessment of their relative match. The basic requirement for equivalence of original and translation in his model is that the translation should have a function (consisting of an ideational and an interpersonal functional component, in the Hallidayan sense) which is equivalent to that of the original. The translation should also employ the equivalent pragmatic means for achieving that function.

As Munday (2001) believes , House translation quality assessment model is based on comparative analysis of the source and target texts. The comparison leads to evaluation of the quality of the translation by focusing and finding `mismatches` and `errors`. In this model, the basic requirement for the equivalence is that should have a function which is corresponding to that of the source text.

House (2015) based on translation studies during past decades such as interdisciplinary studies, cognitive studies ,cross-cultural ,intercultural and translation evaluation defined translation : Translation can be defined as the result of a linguistic-textual operation in which a text in one language is re-contextualized in another language. As a linguantly influenced by a variety of extra-linguististic-textual operation, translation is however subject to and substantially influenced by a variety of extra-linguistic factors and conditions. It is this interaction between `inner` linguistic-textual operation, and `outer` extra-linguistic, contextual factors that makes translation such a complex phenomenon. Some of the interacting factors we need to consider when looking at translation are:

- The structural characteristics, the expressive potential and the constrains of the constrains of the two languages involved in translation;
- The extra-linguistic world which is `cut up` in different ways by source and target languages;
- The source text with its linguistic-stylistic-aesthetic features that belong to the norms of usage holding in the source lingua-cultural community;
- The linguistic-stylistic-aesthetic norms of the target lingua-cultural community;
- The target language norms internalized by the translator;
- Intertextuality governing the totality of the text in the targer culture;
- Traditions, principles, histories and ideologies of translation holding in the target lingua-cultural community;

- The translational `brief` given to the translator by the person(s) or institution commission the translation;
- The translator`s workplace conditions;
- The translator`s knowledge, expertise, ethical stance and attitudinal profiles of the translator as well as their subjective theories of translation." (p.2-3)

3.2 Rubric in Translation Assessment

The first step in translation assessment is to establish a model of quality and then to transform it into a set of metrics that measure each of the elements of that quality. According to Angelelli and Jacobson (2009,p.38): Rubrics allow for a more systematic and holistic grading .A rubric generally contains all sub-components that constitute the construct. It provides descriptive statement of behaviors that candidates may exhibit in a particular sub-component. Since a scoring rubric can be used to holistically score virtually any product or performance, it makes sense to discuss its feasibility for scoring translation. A rubric is developed by identifying the characteristics of translation competence, the primary traits of the product or performance, (i.e. micro-linguistic competence, textual competence, pragmatic competence, strategic competence, etc.) and then delineating criteria used to discriminate various levels of performance as was done earlier with each of these-sub-components.

Muzii (2007) believes that a comprehensive set of criteria must assess the quality of translation from several perspectives during the production process. Making a single metric would not reveal all problems. Creating multiple metrics that assess the various aspect of what is to be measured, in this case translation, can help recompose the overall framework and give an indication of which parts of a process work well and which part does not. Hence,

a reliable and valid rubric in translation assessment would be required to address the aforementioned issues Stevens and Levi (2004) write " At its most basic, a rubric is a scoring tool that lays out the specific expectations for an assignment. Rubrics provide detailed descriptions for what constitutes acceptable and unacceptable levels of performances"(p.3). Riazi (2003) describes the rubric as an attempt to delineate consistent assessment criteria. He emphasizes that it allows teachers and students alike to assess criteria, which are complex and subjective and provide ground for self-evaluation, reflection, and peer review.

There are of course various rubrics in the literature of translation studies. The earlier ones were based on error analysis, examples of which include the one designed by Sager (1983) which was based on three types of errors of linguistic, semantic, and pragmatic elements in five levels for scoring, and the one designed by Mason (1997) in four levels of errors for every single sentence. Rubrics developed later, however, focused on positive points in addition to penalties, similar to the one developed by Rico Perez (2002) which was designed to detect six types of errors and three degrees of seriousness. Later rubrics by Farahzad (1992),Waddington (2001), Sainz (1992), Beeby(2000), and Goff-Kfouri (2005) , were developed of which I mention some of them here.

Frahzad's Rubric

According to Frahzad (1992,p.274), Farahzad checks two features in scoring for each unit of translation ,`a sentence`, in her rubric:

- Accuracy: the translation should convey the information in the source text (ST) precisely, i.e. the translation should be close to the ST norms.

- Appropriateness: the sentences should sound fluent and native, and should be correct in terms of structure.

In terms of scoring, she determined half of score for unnatural translation which convey the ST's meaning, whereas inaccurate translations receive no score, no matter how appropriate and natural the target texts sound. In error recognition items, one score is given for spotting the error and another one for correcting it. Farahzad (1992) introduced two different ways for scoring a long text.

A. It can be scored holistically :since the items assesses a wide variety of competencies, the examiner may find it convenient to approach the text as the unit of translation; the examiner then allots points to each important factor, and the total rating then constitutes the score.

The examiner may, for instance, come up with the following scheme:

1. Accuracy
   20 percent
2. Appropriateness
   20 percent
3. Naturalness
   20 percent
4. Cohesion
   20 percent
5. Style of discourse/choice of words
   20 percent.

According to Farahzad (ibid,pp.276-277) the second method of scoring is as following:

B. The target text must be read two times, first to check the accuracy and appropriateness, the sentence and clause should be the unit of translation. Thus each verb in the source language text marks a score, since verb is the major marker of sentence. Whenever the source language sentence consists of a main clause plus subordinate clause(s), the main clause receives one score and each sub-clause another score. Sometimes the score sentences contain embedded clause. If it is expanded in the process of translation and rendered into a main and a sub-clause, the target version receivers only one score, so as not to reward over translation. Then cohesion and style are checked in each sentence and clause. A correct sentence which does not preserve the content, receives no score. If the target version conveys the message, but in a structure distorts the meaning, the translation receives no score. And if the message is conveyed, albeit in a grammatically unnatural form, the translation receives half of a score.

She (ibid.) adds cohesion and style cannot be checked and scored at the sentence and clause level. The elements of cohesion (e.g. translational, appropriate use of pronouns, linkages, etc.) are spread all over text as are the elements which from the style of discourse (choice of words, grammatical structures, etc.); in checking and others for style, depending on the importance attached to each. If, for instance, the source text is fairly neutral, one may allot a smaller number of points to it than in other cases where the preservation of style is important.

Beeby's Rubric

Allison Beeby (2000,p.189) proposed a model for evaluation experienced in 1996 and 1997. In her model, the exam was marked out of 20:

- 10 points given for 10 specific translation problems that had been selected from the text
- 10 points given for language, and marks were subtracted for grammar mistakes:1 point for incorrect syntax, tense, agreement and word order; ½ a point for incorrect articles, prepositions, and spelling

According to Khanmohammad and Osanloo (2009): the 10 points selected for measuring translation competence were : point1   for the translation of the headline and a full mark for an eye-catching headline. Point 2 for typographical differences, for example capital letters in English and not in other languages. Point 3 to 6 for the translation of three long sentences which involved transfer competence necessary in changing the tense, discourse competence necessary for splitting the sentence up and adding the information to different parts of the text to fit in with coherence expectations in English , and knowledge of syntactic differences between the two languages and understanding the function of the parenthesis in the ST. Point 7 for relevance based on transfer competence, awareness of the TL readers` knowledge and making implicit information explicit where necessary. Point 8 for lexical errors; points 9 for cultural transfer; and point 10 for extralinguistic knowledge (beeby 2000, pp. 192-5).

Goff-Kfouri's Rubric

Goff-Kfouri (2005) points out there have been many suggestions made as to how to mark a translation. Certainly the type of translation whether technical or literary plays a crucial role in the type of correction you choose. The corrector also plays an important role. Some emphasize certain criteria above others. Goff-Kfouri (ibid) declares that there are basically three options an instructor can choose from when correcting a translation.

Although some experienced raters are able to differentiate between a paper that is a 62/100 rather than a 67, for example, a general impression mark is not very beneficial to the student for it does not, in general, provide the reasons for the missing marks

A simple error count is not recommended as a method of marking a student's translation since it rarely gives points for content and does not take into consideration the seriousness of the errors.

Heaton (1990,p. 110) proposed an analytical grid for language courses. However, it can be easily adopted for a translation correction. An analytical grid allows the instructor and raters to set clear criteria for correction based on simple arithmetic

Table 2.3.Goff-Kfouri's rubric

Correction Criteria	5	4	3	2	1
Fluency/Flow					
Grammar					
Terminology					
General Content					
Mechanics	x	X			

The correction criteria are Fluency /Flow, Grammar, Terminology, General Content, Mechanics. The range of scores for each criterion is 1 to 5 except Mechanics for which the maximum score is 3,and the total score would be 23.

Waddington's Rubric

Waddington (2001), points to two themes considering almost all the descriptive or theoretical contributions to translation quality assessment:
(i) Establishing the criteria for a "good translation"
(ii) The nature of translation errors:
- Defining the nature of translation errors as opposed to language errors
- Drawing up a catalogue of possible translation errors
- Establishing the relative, as opposed to absolute, nature of translation errors
- The need to assess quality at both the linguistic the pragmatic levels
(iii) Basing quality assessment on text linguistic analysis

(iv) Establishing various textual levels on a hierarchical basis and linking the importance of mistakes to these levels

(v) Assessment based on the psycholinguistic theory of "scenes and frames"

Furthermore, Waddington (2001) introduced four methods of assessment in his model of translation quality assessment: Method A is taken from Hurtado Albir (1995); it is based on error analysis and possible mistakes are grouped under the following headings:

(i) Inappropriate renderings which affect the understanding of the source text; these are divided into eight categories: *contresens*, *faux sens*, *nonsens*, addition, omission, unresolved extralinguistic references, loss of meaning, and inappropriate linguistic variation (register, style, dialect, etc.).

(ii) Inappropriate renderings, which affect expression in the target language; these are divided into five categories: spelling, grammar, lexical items, text and style.

(iii) Inadequate renderings which affect the transmission of either the main function or secondary functions of the source text.

Method B is also based on error analysis and was designed to take into account the negative effect of errors on the overall quality of the translations. The corrector first has to determine whether each mistake is a translation mistake or just a language mistake; this is done by deciding whether or not the mistake affects the transfer of meaning from the source to the target text: if it does not, it is a language error (and is penalised with – 1 point); if it does, it is a translation error (and is penalised with –2

points). However, in the case of translation errors, the corrector has to judge the importance of the negative effect that each one of these errors has on the translation, taking into consideration the objective and the target reader specified in the instructions to the candidates in the exam paper.

Waddington (2001) considers method C as a holistic assessment method and treats the translation competence as a whole, but requires the corrector to consider three different aspects of the student's performance.

The forth method, Method D is the combination of error analysis method B and holistic Method C in a proportion of 70/30;in other words, method B accounts for 70% of the total result and method C for the remaining 30%. Scoring procedure of each method is demonstrated in details in Waddington (2001) paper.

4. Towards defining an assessment text as a valid instrument in translation quality assessment research

As a result of what was mentioned rubrics help the researcher to provide required numerical data for evaluators to assess the quality of translation. However, it seems the selected text for translation quality assessment needs to have some criteria to be more efficient or considered as a standard evaluation instrument in studies.

As the first characteristic it seems the selected text should be divided into three sections including *Opening*, *center, wrap up as* it is considered by Conde (2009), although he specially focused on evaluation of the text and evaluators. According to Munoz-Martin (2009) Conde (2009) in his study examined: that when studying translation quality

,examining the behavior of evaluators as reflected in their work might be as informative as analyzing the process of translating through in their work might be as informative as analyzing the process of translating through log files to infer regularities in translators` behaviours.

Munoz-Martin (ibid) added Conde (2009) in his thesis examined: the marks of 88 subjects (40 potential readers,25 advanced translation trainees,13 professional translators, and 10 translation teachers working at Spanish or Mexican universities),who worked on four sets of 12 translations each from English into Spanish, two of them on politics and the other two on technological processes of painting machinery. The number of actions did *not* correlate with the final judgment on the quality of a translation, and evaluators showed statistically significant differences in the number of action on the texts, which did *not* parallel differences in judgment. Furthermore, the number of actions lowered constantly in all cases from text 1 to text 48, and this, again, did not have any effect on the judgment either. Marks seemed to concentrate progressively on more salient phenomena, which seemed to weight more on their evaluation. Texts were divided for analysis into three sections of roughly the same length (opening, central, and wrap-up), and Conde found that the correlation between the judgment and the central section was tighter than with the opening section and the wrap up. Our tentative explanation is that the evaluators underwent a learning process when confronted with sets of translations from the same original, and that they learned to concentrate on certain phenomena as their task progress. Also, subjects disregarded some phenomena in

opening sections, probably because they used those sections to contextualize their activity, an influence of their usual behavior as regular reads. On the other hand, when they reached the wrap-up section, they probably had already a notion, if vague, of the quality of the translation.

Therefore ,base on Conde`s findings about the three sections it can be inferred that the standard evaluation text as a translation quality assessment instrument should be a text in a *length* in which there is enough information in its Opening section for the reader to provide her the required condition to *contextualize*. The subjects even could not be wanted to translate the Opening section since both the subjects and evaluators do not probably deal with a valid and appropriate quality of their works in this section. The researcher should consider the role of The Opening part as the mean of contextualizing for the subject since the role of context in comprehending and translating the text is prominent.

Also , as Conde (2011) clamed in his paper some subjects left their work incomplete and as my mates and I while doing our thesis research experienced that some subjects left the last part (i.e. wrap up section of Conde division scale) untranslated, the central section of the text has the special importance.

Another important point is that the readability of the text should be measured by the researcher. In this way the researcher can have comprehensive information about the difficulty level of the text and even its sections (in some cases it might include some paragraphs) ; and knowing about the difficulty level of the text the researcher is more accurate to select the appropriate subjects in order

to inquire the quality of translation. Besides when evaluators assess the text based on models or rubrics, the achieved detailed information by accounting readability statistics, can help evaluators in understanding and finding the *errors* the subjects have done

In addition I believe the subjects should have the required dictionary to have the meaning of the text`s words as the primitive material while translating .In this way the quality of transition can be more specifically concerned in the translation quality assessment studies.

5. Conclusion

To sum up what was discussed in details, a written text plays the main role in translation which is defined as the transformation process of the meaning from source text to the target text. In today translation studies, a text is counted as one of the most basic instruments of the research, specially experimental - researches and since there is constant attempt to make researches more scientific , the author emphasized on the requirement of a standard text or endeavor to standardizing the text as one of the scientific research instruments . In this respect the author considered the researches such as the one done by Conde (2009) focusing on translation quality assessment as one of the most proponent and well-used variables in translation studies and by evaluators . Based on the findings and consideration of the problems translators dealing with while participating in the researches such as limited translation time and the research conditions, researchers should considers some criteria to select a text as an evaluation instrument of translation quality and ability of the participants. The criteria such as the *readability information* of the text, the appropriate *length* of the text in the way, the translator can identify and acquire the *context* are important. Even though providing the required dictionary for participants while rendering would be an efficient step for researchers to achieve the participants` translation qualification. However there is more requirement for researches in this area .

Reference

Al-Qinai, J. (2000) . Translation Quality Assessment. Strategies, Parameters and Procedures ' *Meta, XLV, 3, 2000.*

Angelelli, V. C., & Jacobson, E. H. (2009). *Testing and Assessment in Translation and Interpreting Studies: A Call For Dialogue between research and Practice.* John Benjamins,B.V.

Baker, M.& Saldanha,G .(Eds.)(2009). *Routledge Encyclopedia of Translation Studies.* Newyork: Routledge.

Beeby, A. (2000). *Teaching Translation from Spanish to English.* Ottawa: University of Ottawa Press.

Conde-Ruano, T. J. (2009). It doesn't look bad: Behavior and results of students at evaluating translations. Unpublished monitored research. University of Granada.

Conde-Ruano, T. J. (2011). Translation Evaluation on The Surface of the Texts: a Preliminary Analysis.*The Journal of Specialized Translation,8(15),69-86*

D`Hulst, L. (1996). Unité et diversité de la réflexion traductologique en France (1722-1789), in M. Ballard & L. D'Hulst, *La traduction en France à l'âge classique, Villeneuve d'Ascq* (pp.83-100): Presses universitaires du Septentrion .

Farahzad, F. (1992). Testing achievement in translation classes. In C. Dollerup & A. Loddegaard (Eds.), *Teaching translation and interpreting: training, talent, and experience* (pp. 271-278). Amsterdam/Philadelphia: John Benjamins.

Gharacheh, B. (2005). *The effect of back translation on the assessment of*

translation. Unpublished, Master. thesis, Tehran Azad University.

Goff-Kfouri, C. A. (2005). Testing and evaluation in the translation classroom. *Translation Journal*, 9(2),75-99.

House, J. (1977). *A model for Translation Quality Assessment*, Tübingen: Gunter Narr.

House, J. (1997). *Translation Quality Assessment: A Model Revisited*, Tübingen: Gunter Narr.

House, J. (2006). Text and context in translation. *Journal of Pragmatics*, *38*(3), 323-337.

Hurtado Albir, A. (1995). "La didactica de la traduccion.Evolucion y actual," *X Perspectivas de la Traduccion*, (P.Fernandez, ed.),Valladolid, Universidad de Valladolid,pp.49-74.

Khanmohamad, H .,& Osanloo, M.(2009). Moving toward Objective Scoring: A Rubric for Translation Assessment. *Jounal of English Language Studies*,1(1),131-153.

Masson, I. A. (1997). Pragmatic comutation. In J.Harland (Ed.), *Proccedings of the Australian Theory Symposium Cats` 97* (pp.103-112). Retrieved on September 17,2010, from http://mcs.une.edu.au/~iam/Data/papers/97cats.ps

Munoz Martin.R. (2009) . Expertise and Environment in Translation *Mutatis Mutandis*. 2(1).24 – 37.

Muzii, L. (2007). Adaptation, translation, and multimediality. *Online Translation Journal*,1(4),46-58.

Richards, J. C., Platt, J., & Platt, H. (1992). *Dictionary of language teaching and applied linguistics* (2 ed.).Harlow :Longman .

Rico-Perez, C. (2002). Translation and project management. *Translation Journal*,6(4), 38-52.

Sager, J. C. (1983). *Quality and standards: The evaluation of translations.* In C. Picken (Ed),*The translator`s handbook* (pp. 91-102). London: ASLIB.

Sainz, J .M. (1992). *Student –centered correction of translation*. Amsterdam/ Philadelphia: John Benjam

Stevens, D.D., & Levi ,A. (2004*). Introduction to Rubrics: An assessment tool to save grading time, convey effective feedback, and promote student learning. Stylus Pub LIC.* Retrieved on November 5,2010, from http://catalog.ebay.ca/introduction-to-rubricsins.

Waddington, C. (2001). Different methods of evaluating student translation: The question of validity . *Meta: Translator`s Journal*,46(2),311-325.