

SAL 26/2015

A Summative
Translation Quality
Assessment Model
for Undergraduate
Student
Translations:
Objectivity Versus
Manageability

Received 01/2015

Accepted 05/2015

A Summative Translation Quality Assessment Model for Undergraduate Student Translations: Objectivity Versus Manageability

Suminis studentų vertimo kokybės
įvertinimo modelis: objektyvumas
ir gebėjimas

TRANSLATION / VERTIMAS

Gholamreza Medadian

PhD student of Translation Studies at the University of Isfahan, Iran.

Dariush Nezhadansari Mahabadi

Assistant professor of language teaching at the University of Isfahan, Iran.



<http://dx.doi.org/10.5755/j01.sal.0.26.12421>

Abstract



Research Journal
Studies about Languages

No. 26/2015

ISSN 1648-2824 (print)

ISSN 2029-7203 (online)

pp. 40-54

DOI 10.5755/j01.sal.0.26.12421

© Kaunas University of
Technology 2015

Various scholars from different schools of thought have proposed criteria and/or models or translation assessment. Surprisingly, almost none of them are tailor-made for a manageable summative evaluation of student translation. That is why most translation teachers still draw on holistic and traditional methods of translation evaluation in their exams. These methods are either too holistic or too detailed (and complex) for translation evaluation purposes in educational settings. The holistic approaches that verge on subjectivity are quite manageable for a teacher who is to evaluate of a score of students, whereas the detailed and quantitative models, which are highly demanding on the limited resources of a classroom teacher, are considered highly objective. Feeling the need for a model, which is both manageable and objective, this study aims at reaching a compromise between the subjectivity and the complexity of these approaches

to translation evaluation. Our proposed model draws on and combines the five linguistic equivalences introduced by Koller (1979) and the five-leveled holistic scheme for translation evaluation proposed by Waddington (2001).

Key words:

Translation quality assessment, summative evaluation, equivalence, student translation.

According to the early theories of translation, translation evaluation was performed quite unsystematically by means of a single master criterion such as *the principle of equivalent response* (Nida, 1964) or binary criteria like word for word, sense for sense or faithful/ unfaithful translations. The call for systematic and objective translation evaluation goes back to the 1959 FIT conference where the theme was translation quality. From then on, there has been much progress in translation quality assessment in order to come up with more systematic and objective assessment scenarios. However, there is not much consensus among the scholars and practitioners on the assessment criteria. As Hatim and Mason (1997, p.199) rightly observe this wide area of research still deserves more attention and is quite under-explored.

According to Martinez Melis (1997 and 2001) there are, basically, three sub-categories with translation quality assessment as a whole: the assessment of published translations (principally religious and literary works), the assessment of the professional translators' competence and assessment of student translations in educational contexts. From among these areas, it is the assessment of student translation that has been mainly neglected.

Inevitably, various types of translation evaluation entail different evaluation criteria. According to Hurtado Albir (1990), the criteria for the evaluation of published literary and religious texts have been faithfulness to the original and the subjective, mostly aesthetic, quality of the translated text. More often than not such an evaluation has been mainly performed by non-academic people without due attention to the source texts. Naturally, such evaluations have been very subjective. Conversely, within academic circles, various criteria have been suggested to evaluate published translations (e.g., Larose, 1989; Nord, 1991; House, 1997/2009, and Williams, 2001 and 2004). Professional translator's competence has been mostly gauged through quantitative *correcting* and *grading* scales. Correcting scales provide the evaluator with an error typology while the grading scales allocate scores to the error types detected in translations. For example, Gouadec (1981 and 1989) proposes one of the most comprehensive correcting scale including 675 lexical and syntactic error types to assess professional translator's competence. Such scales are usually composed of numerous parameters and claim absolute objectivity.

The most under-researched area in evaluation is the student translations. The models and scales proposed for the evaluations of published translations and professional translators have dramatically overshadowed this area through the years for some false reasons. Firstly, it seems that student translations are not considered as valuable and influential as the literary works and the students are considered inferior to the professional institutional translators. Secondly, it

Introduction

is assumed that the assessment models proposed for the published translations and the quantitative scales developed for assessing professional translator's competence can be readily adopted and applied for pedagogical purposes. However, applying such models, which may contain over a hundred parameters, is not a viable option for a classroom teacher who is supposed to correct a pile of examination papers in a limited time at the end of the term. Without evaluation and assessment students trainees will not know how to enhance the quality of their translations systematically (Honig, 1997) and the teachers will not have reliable a road map to allocate objective scores to their students. If translation trainees do not improve themselves, we may not have high quality literary translations and high profile professionals for whom we can propose translation quality assessment models.

Some questions come into mind regarding the adoptability of translation quality assessment models proposed for the literary works and professional translators in the educational settings. Does teacher have enough resources (time and energy) to score student translations objectively using the criteria set for professional translators? Is it fair to expect translation trainees to match senior experienced translators' competence? Can we assess the quality of student translations without considering the source texts (as non-academic people do in evaluating the quality of literary and religious works)? The answer to the above questions cannot be a 'yes'.

In a study carried out by Waddington (2001), a questionnaire was distributed among European and Canadian translation teachers to find out about the assessment methods they employ. It was found that 38.5 % of the respondents still employ a holistic (subjective) method to judge the student translations and to the author's surprise, they base their judgment on the requirements of the professional translators. Thus, there is a need to develop a new practical model to assess the student translations based on the requirements of the translation training programs. Waddington (2001) notes that teachers use three methods to evaluate student translations: error analysis, holistic and a combination of error analysis and holistic judgment. He believes that the sum of errors may not directly reflect the quality of translation and as Pym (1993) notes the distinction is not between black and white, but between various levels of grey. Teachers still employ the traditional and subjective criteria to assess student translations because the supposedly objective criteria proposed for evaluation is not manageable in educational contexts.

Feeling the need for further research in the area of evaluating translation trainees, Martinez Melis (1997, p.156) outlines various objects, types, functions, aims and means of student translation evaluation:

- *Objects*: student translator competence, study plans, programs
- *Types*: product assessment, process assessment, qualitative assessment
- *Functions*: diagnostic, summative, formative
- *Aims*: academic, pedagogical, speculative
- *Means*: translations, evaluation criteria, correcting criteria, grading scales, tests, exercises, questionnaires, etc.

As to the limited scope of this paper, the authors only focus on the product-based summative evaluation of student translations in undergraduate programs. The organization of the paper is as follows. Firstly, the paper scrutinizes the merits and demerits of the various theories and models of translation to see their potential as the basis for a manageable but objective evaluation of student translations. Then, we propose our model that includes a correcting and a grading scale for the summative evaluation of student translations.

In order to develop an evaluation scenario a road map is needed. Such a map can be provided by various translation theories. In fact, one of the important objectives (if not the most important) of translation theorization has been the concern for telling high quality translations from the low quality ones. According to House (2009b, p.222) every “translation quality assessment presupposes a theory of translation”. Different theories offer different quality yard sticks for translation evaluation purposes.

Pre-systematic views

Early translators use mainly dichotomous criteria such as word-for-word versus sense-for-sense, literal versus free, and faithful versus unfaithful to tell ‘good’ translations from the ‘bad’ ones’, but can a translation be absolutely good or bad, faithful or unfaithful? Some emphasized the preservation of some abstract constructs such as ‘spirit’ and ‘truth’ to assess translations and translators (Kelly 1979, p.205). According to Munday (2008, p.24) some even considered concepts like “the creative energy of a text” as the cornerstone of an adequate translation. The early theorists mainly used a metaphorical language to describe and explain translation. Some theorists proposed few parameters for evaluation. For example, Larson (1984) in the *meaning-based translation*, proposed only three parameters of *accuracy*, *naturalness* and *faithfulness* as the determining factors of a good translation, but she did not operationalize these criteria properly. Munday (2008, p.26) mentions an early theorist like Dryden who expressed his criteria for a translation via a few reductionist rules, but did not follow those criteria even in his own practice. The reason for such a deviation from one’s own prescription could be nothing but the subjectivity of proposed rules.

Post-modernist views

Postmodern views on translation strived to challenge the existence of a stable meaning in a text and believed in the relativity of language. As a matter of fact, subjective and relative criteria for ‘adequate’ translation are legion in postmodern translation theories. For example, Pound (1918/2004) and Benjamin (1969/2004) both advocated highly innovative and experimental translations as ‘good’. Benjamin called for highly literal translations in order to let the foreign text ‘shine’ through in translation. He believed that a ‘pure language’ would be revealed as a result of direct interaction between the two languages. Stolze (1992) states that a ‘good’ translation comes into existence, when a translator is able to identify fully with the source text. However, he does not specify any means by which one can objectively measure the amount of translator’s identification with the source text.

Various Approaches to Translation Evaluation

According to the postmodernists, readers may potentially have infinite interpretations of the same original text. This fuzzy definition of meaning and interpretation of source text, though attractive and thought provoking, is something that defies objectivity as one of the most desirable features of an evaluation scenario in educational context where the students are compared with each others in terms of their scores. Another main drawback of postmodern views is that they do not put the criteria they prescribe into measurable terms and any evaluation scenario arising from them lacks objectivity.

Semi-systematic response-based views

The early systematic views stemmed from the leading linguistic theories like Structuralism and Transformational Generative Grammar. Taking a behaviorist stance, Nida (1964) considered rebuilding the “equivalent response” on TT (target text) readers as the master key of ‘good’ translation. Twenty years later, Nida and Taber (1982) attempted to operationalize the notion of equivalent response through cloze tests and parallel translations to lend validity and reliability to their views. They believed that a translator should opt for the “closest natural equivalent” in the target text to achieve “dynamic equivalence” (1964, p.166).

This response-based view received severe criticisms. Lefevere (1993, p.7) criticized Nida for sticking to sentence level and disregarding extra-sentential levels of analysis. Larose (1989), Van Den Broeck (1978) and Hu (1992; 1993) declared equivalent response as totally ‘impossible’. Hu argues that equivalent response entails numerous translations of the same original for different readers, because every reader may have a unique response to a certain translation. Newmark (1988) tried to tackle the criticisms raised against Nida by viewing the equivalent response as non-working out of ST’s (source texts) time and space, but he introduced more subjective notions like translation as an ‘art’ or ‘craft’ that leave large room for subjectivity quality assessment.

There are some key limitations in the response-based views of translation evaluation. Firstly, they are highly target text and target reader oriented; the intrinsic values of the ST are largely disregarded. Secondly, the immeasurability of the ST and TT readers’ responses turn equivalent response into an indefinitely vague and subjective criterion. In addition, employing only one single master criterion (equivalent response) to evaluate a complicated phenomenon like translation is a very reductionist, indeed. Having these limitations, equivalent response cannot be a reliable foundation for an objective evaluation of student translations.

Functionalist and skopos-based views

The early functionalist approaches focus only on one master criterion when explaining the translation phenomenon. According to Reiss (1971/2000), text types and their corresponding functions are the most influential factors in determining translation strategies. Subsequently, Reiss and Vermeer (1984) considered the dominant function (skopos) of the text as the criterion against which the quality of a translation can be measured. To them all the decisions resorted to by the translators are subordinate to the primary purpose of the target text, but they did not propose any procedures for translation evaluation. In skopos

theory's target text oriented framework, the purpose of target text justifies the tools (translation strategies) whatever they may be.

Working within the functionalist paradigm, Nord (1991) considers translation as a communicative act and believes that in translation evaluation micro-textual error analysis may not be a reliable tool if applied alone, because a translator may omit or add something to the translation to meet the requirements of the target text's macro-level function. She tries to devise an overall text analysis model, but most of the times she considers only one specific parameter in her analysis and maintains that it is not possible to have a comprehensive evaluation of a translation.

The formalist polysystem theory (Even-Zohar, 1978/2004) investigates translation and its quality in isolation from its source to determine the status and the function of a translation in the target literary system. However, it does not provide a scenario for translation evaluation. This theory considers translation and its function as an independent and stand alone work quite like an original. Poly-system theory largely disregards the intrinsic value and the status of ST in itself. Borrowing ideas from polysystem theory, theorists like Toury (2012) and Chesterman (1997) try to establish various translation norms and standards based on the observation of the regularity of the behavior in the translated texts and the translators, but they fall short of proposing any practical procedures for translation evaluation. To them adherence to the prevalent norms of translation in the target text is the corner stone of 'good' translation. Naturally, a flexible notion like norm entails a lot of subjectivity on the part of the translation evaluator. If the norms of translation were already established for a specific society in specific time, we could use them as criteria to judge the quality of any student translation.

The main concern of functionalist and polysystemic views of translation is the evaluation of published (literary and religious) translations. Either they do not offer any procedures for translation evaluation or offer highly complicated procedures for textual analysis. In addition, much like response-based theories, they are highly target-text oriented. Following skopos-based views on evaluation, every student translation could be judged as faultless resorting to the excuse that the translation is serving a certain purpose (skopos), which might be different from the original purpose of the ST. These major limitations defy the adoptability of these views as the base for correcting and grading scales for the evaluation of student translations.

Complex and multi-parameter views

These views mainly draw on the different dimensions of language introduced and explained by theories such as (critical) discourse analysis, speech act, pragmatics, argumentation theory, etc. At times, eclectic mixtures are used to compare and contrast STs and TTs and make quality statements. The distinguishing feature of these views is the consideration of both ST and TT in the real life contexts and analyzing the texts on supra-sentential levels.

House (1997 and 2009a), borrowing ideas from theories such as speech act and pragmatics developed a highly complicated model for translation quality assessment. Although working within a functionalist framework, she criticizes

the target-oriented nature of the previous functional theories of translation, and maintains a more balanced appreciation of the roles of ST and TT in her model. According to House (1997, pp.66–69) ST and TT must be analyzed from various dimensions (semantically, pragmatically, textually, etc.) to look for a relative “functional equivalence”. She distinguishes between two basic types of translations. Where an overt translation flavors, symbolically, like a translation and the readers know that they are reading a translation, a covertly translated text seems like an original in the target culture because of being passed through a “cultural filter”. Cultural filter is a highly vague construct and entails a subjective judgment that is far from desirable for a translation quality assessment model that claims objectivity. For House, the word ‘culture’ encompasses textual and linguistic norms and conventions embedded in every linguistic community. She admits that “translation is at its core a linguistic-textual phenomenon” and it must be evaluated as such. As a matter of fact, House (2001, p.254) advocates a combination of objective linguistic evaluation of translation (which is inter-subjectively verifiable) plus a holistic value judgment taking into account the macro-level (social and cultural) qualities of ST/ TT.

Although House’s model (1997) has tackled many of the criticisms leveled against the previous views, it is highly impractical for real-life educational purposes. If we consider the previous theories and their promises for quality assessment as highly subjective and reductionist, House’s model is objective, but far too unmanageable for educational purposes. House (1997; p.64) admits that it sounds “unlikely that translation quality assessment can ever be objectified in the manner of natural science”. That is why she incorporates cultural and social loads of ST and TT in her analysis. This is something that must be considered in every evaluation scenario. House’s model is mainly used for corpus-based studies within the *covert translation project* in the University of Hamburg. Its adoptability for educational settings with teacher’s limited resources does not seem viable.

Al-Qinai (2000) attempted to develop an ‘empirical’ framework for quality assessment drawing on some parameters like textual typology, formal correspondence, thematic coherence, reference cohesion, lexical-syntactical properties, etc. Such an eclectic model only adds to the number of the parameters involved in the assessment process, something that is to be avoided in order to have a manageable evaluation criterion for student translations.

Williams (2004, pp.21–30) draws on argumentation theory, which is one aspect of discourse analysis, and *New Rhetorics* to build a general and whole-text framework for non-literary (instrumental) translation evaluation. He argues that argumentation is present in every text ranging from the most scientific to the most literary and thus his proposed model can be applied to all text types. Williams (2004, pp.32–65) explains that every argument is composed of six components (claim(s), ground(s), warrant(s), backing(s), qualifier(s), and rebuttal) which can be compared cross linguistically. The model checks whether the translations render these components accurately. Later on, Williams turns into eclecticism and borrows ideas from various scholars and under the designation ‘rhetorical typology’ adds more parameters like organizational relations, conjunctives, inference indicators, propositional functions, arguments type, figures of speech and nar-

rative strategy to his “multiple-parameter grid” for translation evaluation. There are serious limitations if one wants to adopt William’s model to evaluate student translations. Firstly, identifying these components does not seem to be an easy task at all and not every piece of text includes all these components. Secondly, the ‘accuracy’ of rendering each parameter is not objectively measurable.

The multi-parameter and eclectic models may be of great academic interest and help researchers understand the translation phenomenon and evaluate published translations or professional translators’ competence but they cannot be applied in evaluating student without considering the specific requirements of the educational context. It seems that the authors of these models are under the influence of eclecticism as an ‘enlightened’ method which has strong supports in second language learning literature (see Brown, 2007, and Swain&Lapkin, 1995). However, if eclecticism has worked for second language teaching, this does not justify its application in translation evaluation.

Trying to explore the advantages and disadvantages of the existing translational evaluation scenario, Williams (2004) divides all the existing models into two broad categories: *qualitative* and *quantitative*. Quantitative models are mainly developed and used by organizations and bureaus for certification and examination purposes in professional contexts. They mainly focus on a small random sample of a text and conduct sub-sentence error analysis and error counts to score translations based on predefined *correcting* and *grading* scales. Use of short sample of professional translator’s translation is quite similar to summative evaluation of student translations where a usually short text is translated and then evaluated in the final exams.

One of the important factors considered in translation evaluation in various models is *error gravity*. According to Nord (1996) the higher the level of errors, the stronger its impact on the whole text under evaluation. Thus, to her, errors with the highest gravity are pragmatic errors. They are followed by cultural errors and the least important errors are the linguistic ones. Larose (1989) emphasizes the importance of the context in which the errors occur as determining factor in the seriousness of the error. Some scholars maintain that the degree to which an error violates the effectiveness or the functioning of the target text determines the error gravity (Dancette, 1989 and Hurtado Albir, 1995). According to Martinez Melis and Hurtado Albir (2001) error gravity must be analyzed from functionalist stance to see: the effect of the error on the texts as a whole, its effect on the cohesion and coherence of the target text, the extent of departure from the meaning of the original text, the degree of violation of the target text communicative function (including text type conventions) and its impact on the translation skopos.

The main purpose of evaluation in the educational context is to measure student translator’s skills. According to Stansfield *et al.* (1992), there are two basic translation skills: *accuracy* and *expression*. Accuracy is related to source text processing and expression deals with target text production. Hatim and Mason (1997) also emphasize the inclusion of source text processing skills, transfer skills and target text processing skills into every evaluation scenario.

Alongside error gravity, *error typology* is another important aspect in translation

Requirements of a Practical Model for the Evaluation of Student Translations

evaluation literature. Hurtado Albir (1995) identifies two broad types of errors: errors dealing with the understanding of the source text (language errors) and those dealing with the production or the expression of target text (transfer errors). He also considers the faulty transfer of the primary or the secondary function of the source text as another main source of error in translation. Following this error typology, he defines two gravity levels for errors. For every serious error the translation receives -2 and for a minor error -1. One score is allocated to a good translation solution and +2 for an exceptionally good translation solution. However, there are cases in which one cannot decide between major and minor error. Thus, it seems that more than two levels must be defined for error gravity. To sum up, for a model to work as a guide for the evaluation of student translation, it needs to be:

- *Objective*: the postmodern and response-based views are highly impressionistic and their application in educational context leaves much room for subjective judgments;
- *Practical and manageable*: multi-parameter and complex models consider numerous parameters in their analysis for the sake of objectivity and validity but they are far too unmanageable and complex for a real life summative evaluation of student translations;
- *Non-reductionist*: except multi-parameter models, the other views and models make use of only one master criterion (e.g., skopos or equivalent response) or a binary criterion (e.g., free/ literal or overt/ covert) to evaluate a relatively complex phenomenon. Reductionism is a serious threat to validity of an evaluation model;
- *Tailor-made*: the existing models could not be employed to assess the quality of all three areas introduced by Martinez Melis. Most models focus on the evaluation of a published translation and professional translator's competence. We need a tailor-made model to evaluate student translator's competence summatively.
- *Bi-directional*: most theories and their implications for evaluation are either source or target text oriented. We need a model that considers both language and translation skills.
- *Objective and holistic* at the same time, because the sum of micro-textual errors may not determine the quality of a translation and translation evaluation cannot reach to the level of physical sciences in its accuracy.

Such a model needs a scoring system including:

- A correcting scale that defines the error types hierarchically;
- A grading scale that determines the levels of error gravity hierarchically.

It must be noted that if an evaluation model draws only on one master criterion or a dichotomy, the evaluation task becomes quite manageable, but the translation evaluation will be highly impressionistic. On the other hand, if the model draws on a complex analysis, the objectivity can be guaranteed to some extent, but its application will be time consuming and unmanageable for summative translation evaluation of the student translations. Complex models may be of great academic interest, but they are not used by practicing teachers.

Therefore, a model that fulfills both manageability and objectivity is required for students translation evaluation. Such a model must necessarily draw on a few manageable parameters. In the next section, the above mentioned requirements will be put into the framework of a manageable multi-parameter model for translation evaluation.

As we showed in the previous sections, the existing models for translation evaluation cannot be readily adopted for educational settings where texts are to be translated and evaluated. The existing models either draw on too many parameters in their analysis and are, thus, too unmanageable for the summative evaluation of student translations or they are too reductionist and include only a couple of highly impressionistic criteria and thus verge on extreme subjectivity. According to Pollitt (1991) in judging tests, it is optimistic to claim “five reliable bands” to measure the ability we are observing. Inclusion of five levels in analysis can promise both objectivity and manageability when one is evaluating the translation of a usually small text by student translators in the summative exams. Koller (1979, pp.99–104) introduces exactly five equivalences for translation that could be readily adopted for translation evaluation purposes as levels of correcting scale. Koller’s equivalences are *denotative*, *connotative*, *text-normative*, *pragmatic* and *formal* that are hierarchically ordered as to the level of task difficulty.

Denotative equivalence refers to equivalence of the referential content of source and target texts. This equivalence could be considered as the lowest (the easiest and the least demanding) level of student’s translation competence. This has mainly to do with the second language proficiency of a student translator (language skill). For example, any direct deletion or addition could go under this category. A student that does not know the denotative meaning of a source text item may resort to deletion in most of the cases. *Connotative equivalence* concerns the suitability of translator’s lexical choices and the ability to understand and transfer the connotative meanings between source and target texts. Realizing this equivalence demands higher understanding and production skills than the denotative level (both language and transfer skills). *Text-normative equivalence* has to do with text types and suitability of student’s lexical choices regarding the text type under translation (transfer skill). *Pragmatic equivalence* is the ability to understand and produce pragmatically acceptable equivalences across semiotic borders. This entails familiarity with the pragmatic aspects of source and target texts and the way they correspond to each other (transfer skill). *Formal equivalence* deals with aesthetic properties of the texts including word plays, use of puns and highly expressive features of the ST and TT. This is the most difficult level of equivalence and it is highly demanding on the production skills of the translator trainees. The type of equivalence one is looking for depends largely on the text type under translation.

If we consider these multiple equivalences as the base for our model’s error taxonomy, we will have denotative, connotative, text-normative, pragmatic and formal errors in the model. The order of equivalences from the easiest to the most difficult can serve as the basis for determining error gravity. Since languages

Developing a Model for the Summative Evaluation of Student Translation

agree more on micro than macro level, it is naturally more difficult to find a pragmatic equivalence than a denotative one. Therefore, in the evaluation scenario, the higher the level of equivalence required, the higher the demand on the student translator's skill and, thus, the higher the gravity of error in the grading scale (see Nord, 1996). Thus, there are five main types of errors in the correcting scale and five corresponding levels for the error gravity on the grading scale.

If we assume that the total score is 100, 70 points are allocated to linguistic error analysis based on correcting and grading scales proposed earlier according to Koller's equivalences. Thus, our correcting scale would be as follows:

- Every denotative error deduces 1 point;
- Every connotative score deduces 2 points;
- Every text-normative error deduces 3 points;
- Every pragmatic error deduces 4 points;
- Every formal error deduces 5 points.

Some exceptions will be added to preceding scoring scale, because translation is rarely done in a word-by-word manner.

- An error is only punished for the first time. If repeated no point is deduced;
- If students change something in order to achieve an equivalent in a higher level of language, no point is deduced;
- No scores are deduced if an error is compensated for somewhere else in the translation.

Because translation is a relative activity in nature and translations with equal number of errors may still be of different quality levels (see House, 2001; Waddington, 2001; Pym, 1993), the sum of linguistic translation errors (both language and transfer errors) may not directly indicate the quality of the translation and there is a need to add a holistic parameter to the our evaluation scenario. Therefore, 30 points (out of 100) are allocated to an overall holistic appreciation of the rater about translation quality.

We equip our model with the holistic band proposed by Waddington (2001) to capture the holistic appreciation of raters of student translations (Table 1). The marks allocated to each level are tripled to give the maximum score of 30 (in the original band the maximum possible score is 10 for the a translation rates as *level 1*, while the translation which is marked as *level 5* by a rater receives maximum between 1–2 scores). Waddington's band considers three holistically judged criteria, which are accuracy of transfer of the ST content (language skill), quality of expression in the TL (transfer skill) and degree of task completion. Thus, in our proposed model a translation which satisfies the qualities mentioned for *level-1* receives maximum 30 scores, whereas the one which satisfies *level-5* qualities will receive only 6 scores. As it can be seen, *level-1* translation can receive a score ranging from 24 to 30, which allows a translation rater to distinguish more neatly between the quite similar student translations.

This scenario is, then, applied to a student translation of about 300 words including two texts of about 150 words each (two different texts are chosen in order to increase the reliability of the evaluation). If possible, two raters can judge a student translation in order to increase inter-subjective reliability.

Level	Accuracy of transfer of ST content	Quality of expression of TL	Degree of task completion	Mark(s)
Level 1	Complete transfer of St information; only minor revision is need to reach professional standard	Almost all the translation reads like a piece originally written in TL. There may be minor lexical, grammatical and spelling errors	successful	24–30
Level 2	Almost complete transfer; there may be one or two insignificant inaccuracies; require certain amount of revision to reach professional standard	Large sections read like a piece originally written in TL, but others read like a translation. There are a considerable number of lexical, grammatical or spelling errors	almost completely successful	21–24
Level 3	Transfer of general idea(s), but with a number of lapses in accuracy; needs considerable revision to reach professional standards	Certain parts read like a piece originally written in TL, but others read like a translation. There are considerable number of lexical, grammatical or spelling errors	adequate	15–18
Level 4	Transfer undermined by serious inaccuracies; through revision required to reach professional standard.	Almost the entire text reads like a translation; there are continual lexical, grammatical or spelling errors.	inadequate	9–12
Level 5	Totally inadequate transfer of ST content; the translation is not worth revising.	The candidates reveal a total lack of ability to express themselves adequately in TL	totally inadequate	3–6

Table 1. Description of Five Levels of Holistic Evaluation (adopted from Waddington, 2001, p.22).

A student translation that earns at least 50 out of 100 passes the summative final exam. Scores higher than 80 are marked as *excellent*, ones ranging from 60 to 80 are marked as *good*, and those ranging from 50 to 60 are considered *acceptable*: Those falling below 50 are marked as *unacceptable* translations in this evaluation model.

Various criteria and models have been proposed for translation evaluation. Almost none of them are tailor-made for the summative evaluation of student translation at undergraduate level. They are targeted at evaluating literary works and professional practicing translator's competence. Applying most of these models is quite time consuming and demanding on the classroom teacher's limited resources. That is why most of the teachers still draw on purely subjective methods of translation evaluation in the final exams in under-graduate levels. If a model is to gain popularity among the practicing teachers for translation evaluation, it must enjoy both objectivity and manageability.

We proposed a model that includes five types of equivalences in various linguistic levels as a guideline for a correcting scale and five corresponding error gravities

Conclusion

in the grading scale to judge the quality of student translations quantitatively. As it has been rightly argued by some researchers (e.g. House, 2001; Waddington, 2001, and Pym, 1993), the total sum of linguistic errors cannot directly reflect the quality of translation. Therefore, 70 percent of total scores is determined by error analysis (following Koller, 1979) and the remaining 30 scores are determined by evaluator's holistic appreciation of the quality of translation (Waddington, 2001). The main rationale for choosing this combination was to arrive at a manageable model to evaluate student translation in pedagogical contexts.

Besides employing the model for summative assessments, one can also use this model for formative evaluation of student translation in order to find out the weak and strong points of the student translators during a translation course. It must be remembered that for purposes other than evaluation of student translations, there are other highly reliable and objective models existing in the literature that could be employed for the evaluation of literary and professional translations.

Further research is required to see the model in action and find its workability and compatibility with the various institutional contexts. It is recommended to apply this model to various text types. The validity and reliability of the model could be displayed better in evaluating the quality of the texts that include the maximum number of five equivalences proposed in by Koller (1979). It is believed that this model could be applied to evaluate the translation of all texts types in the educational settings.

References

1. Al-Qinai, J., 2000. Translation Quality Assessment. Strategies, Parameters and Procedures. *Meta: Journal des traducteurs/ Meta: Translators' Journal*, 45(3), pp.497–519.
2. Benjamin, W., 1969 (2004). The Task of the Translator. *In: Venuti, L. (ed.) Translation Studies Reader*. Routledge, pp.75–82.
3. Broeck, R. van den, 1978. The Concept of Equivalence in Translation Theory: Some Critical Reflections. *Literature and Translation*, Leuven: Academic, pp.29–47.
4. Brown, H. D., 2007. Teaching by Principles: An Interactive Approach to Language Pedagogy. Pearson Education.
5. Chesterman, A., 1997. Memes of Translation: The Spread of Ideas in Translation Theory. John Benjamins Publishing. <http://dx.doi.org/10.1075/btl.22>
6. Dancette, J., 1989. La faute de sens en traduction, TTR, 2–2, numéro spécial (L'erreur en traduction), pp.83–99.
7. Even-Zohar, I., 1978 (2004). The Position of Translated Literature within the Literarypolysystem. *In: Venuti, L. (ed.) The Translation Studies Reader*, London-New York: Routledge, 2nd edition, pp.199–204.
8. Gouadec, D., 1981. Paramètres de l'évaluation des traductions. *In: Meta XXVI (2)*, pp.99–116. <http://dx.doi.org/10.7202/002949ar>
9. Gouadec, D., 1989. Comprendre, évaluer, prévenir. *Pratique, enseignement et recherché face à l'erreur et à la faute en traduction*. *In: TTR (Traduction, Terminologie, Rédaction) 2 (2)*, pp.35–54. <http://dx.doi.org/10.7202/037045ar>
10. Hatim, B., Mason, I., 1997. The Translator as Communicator. London-New York: Routledge.
11. Honig, H. G., 1997. Positions, Power and Practice: Functionalist Approaches and Translation Quality Assessment. *Current Issues in Language & Society*, 4(1), pp.6–34. <http://dx.doi.org/10.1080/13520529709615477>

12. House, J., 1997. *Translation Quality Assessment: A Model Revisited*. Tübingen: Gunter Narr.
13. House, J., 2001. *Translation Quality Assessment: Linguistic Description versus Social Evaluation*. *Meta: Journal des traducteurs/ Meta: Translators' Journal*, 46(2).
14. House, J., 2009a. *Translation*. Oxford-New York: Oxford University Press.
15. House, J., 2009b. *Quality*. In: Baker, M., Saldanha, G. (eds.) *Routledge Encyclopedia of Translation Studies* (2nd edition). London-New York: Routledge, p.222.
16. Hu, Q., 1992. *On the Implausibility of Equivalent Response (Part I)*. *Meta*, 37(2), pp.289–301 <http://dx.doi.org/10.7202/003148ar>
17. Hu, Q. (1993). *On the implausibility of equivalent response (part IV)*. *Meta: Journal des traducteurs/ Meta: Translators' Journal*, 38(3), pp.449–67. <http://dx.doi.org/10.7202/003147ar>
18. Hurtado Albir, A., 1990. *La notion de fidélité en traduction*, Paris: Didier Érudition.
19. Hurtado Albir, A., 1995. *La didáctica de la traducción. Evolución y estado actual*. In: Hernandez, P., Bravo, J. M. (dir.), *Perspectivas de la traducción*. Universidad de Valladolid, pp.49–74.
20. Kelly, L. G., 1979. *The True Interpreter*. Oxford: Blackwell, p.205.
21. Koller, W., 1979. *Einführung in die Übersetzungswissenschaft*. Heidelberg-Wiesbaden: Quelle und Meyer.
22. Larose, R., 1989. *Théories contemporaines de la traduction* (2e éd.). Québec: Presses de l'Université du Québec.
23. Larson, M. L., 1984. *Meaning-Based Translation: A Guide to Cross-Language Equivalence*. Lanham, MD: University Press of America.
24. Lefevere, A., 1993. *Translating Literature: Practice and Theory in a Comparative Literature Context*. New York: The Modern Language Association of America, p.7.
25. Martínez Melis, N., 1997. *Évaluation et traduction: Cadre de recherche sur l'évaluation dans la didactique de la traduction (these)*. Universitat Autònoma de Barcelona.
26. Martínez Melis, N., Hurtado Albir, A. (2001). *Assessment in Translation Studies: Research Needs*. *Meta: Journal des traducteurs/ Translators' Journal*, 46(2), pp.272–287. <http://dx.doi.org/10.7202/003624ar>
27. Munday, J., 2008. *Introducing Translation Studies* (2nd edition). London-New York: Routledge, pp.24–26.
28. Newmark, P., 1988. *A Textbook of Translation*. New York-London: Prentice Hall.
29. Nida, E. A. 1964. *Toward a Science of Translating*. Leiden: E. J. Brill.
30. Nida, E. A., Taber, C. A., 1982. *The Theory and Practice of Translation*. Brill Academic Pub.
31. Nord, C., 1991. *Text Analysis in Translation: Theory, Methodology, and Didactic Application of a Model for Translation-Oriented Text Analysis*, Amsterdam-Atlanta: Rodopi.
32. Nord, C., 1996. *El error en la traducción: categorías y evaluación*. In: Hurtado Albir, A. (dir.), *La enseñanza de la traducción*, pp.91–103.
33. Pollitt, A., 1991. *Response to Charles Alderson's Paper 'Bands and Scores'*. In: Alderson, Ch., North, B. (eds.), 1991. *Language Testing in the 1990s*. London: Macmillan, pp.87–94.
34. Pound, E., 1918 (2004). *Guido's Relations*. In: Venuti, L. (ed.) *Translation Studies Reader*. Routledge, pp.86–93.
35. Pym, A., 1993. *Epistemological Problems in Translation and its Teaching*. Calaceite: Ediciones Caminade.
36. Reiss, K., Vermeer, J. H., 1984. *gungeinerallgemeinenTranslationstheorie*. Tübingen: Niemeyer.

37. Reiss, K., 1971 (2000). *Möglichkeiten und Grenzen der Übersetzungskritik*. Munich: M. Hueber [translated by E. F. Rhodes) *Translation Criticism: Potential and Limitations*. Manchester: St Jerome and American Bible Society.
38. Stansfield, C. W., Scott, M. L. and Kenyon, D. M., 1992. The Measurement of Translation Ability. *In: The Modern Language Journal* 76 (IV), pp.455–67. <http://dx.doi.org/10.1111/j.1540-4781.1992.tb05393.x>
39. Stolze, R., 1992. *Hermeneutisches Übersetzen*. Tübingen: Gunter Narr.
40. Swain, M., Lapkin, S., 1995. Problems in Output and the Cognitive Processes they Generate: A Step Towards Second Language Learning, *Applied Linguistics*, 16(3), pp.371–391. <http://dx.doi.org/10.1093/applin/16.3.371>
41. Toury, G., 2012. *Descriptive Translation Studies and Beyond* (Rev. edition). John Benjamins Publishing. <http://dx.doi.org/10.1075/btl.100>
42. Waddington, C., 2001. Should Translations be Assessed Holistically or Through Error Analysis? *Hermes, Journal of Linguistics*, 26 (2001), pp.15–38.
43. Williams, M., 2001. The Application of Argumentation Theory to Translation Quality Assessment. *Meta: Journal des traducteurs/ Translators' Journal*, 46(2), pp.326–344.
44. Williams, M., 2004. *Translation Quality Assessment: An Argumentation-Centered Approach*. University of Ottawa Press, pp.1–85.

Santrauka

Gholamreza Medadian, Dariush Nezhadansari Mahabadi. Suminis studentų vertimo kokybės įvertinimo modelis: objektyvumas ir gebėjimas

Skirtingų mokyklų mokslininkai yra pateikę nemažai vertimo įvertinimo kriterijų ir/ arba modelių. Tačiau beveik nė vienas iš tų modelių ar kriterijų nėra skirti suminiam studentų vertimų įvertinimui. Štai kodėl dauguma vertimo dėstytojų per egzaminą vis dar pasikliauja holistiniais ir tradiciniais vertimo vertinimo modeliais. Šie metodai yra arba per daug holistiniai, arba pernelyg detalūs (ir sudėtingi) vertimo vertinimo tikslams edukaciniuose procesuose. Holistinis požiūris, kuriam būdingas subjektyvumas, yra visiškai tinkamas dėstytojui, kuris vertina pažymiu, tuo tarpu detalūs ir kiekybinis modeliai, kuriems reikia daugiau jau ir taip ribotų dėstytojo resursų, yra laikomi ypač objektyviais. Suprasdami nesudėtingo ir objektyvaus modelio reikalingumą, straipsnio autoriai siekia kompromiso tarp šių požiūrių subjektyvumo ir sudėtingumo. Mūsų siūlomas modelis siekia suderinti penkis lingvistinius ekvivalentus, kuriuos įvedė Kolleris (1979), ir Waddington (2001) pateiktą holistinę penkių lygių schemą vertimui įvertinti.

About the Authors

Gholamreza Medadian

PhD student of Translation Studies at the University of Isfahan, Iran.

Academic interests

translation teaching, translation quality assessment, translation contact.

Address

Pardis Language Institute, Sizdah Khaneh, Buali Street, Hamedan, Iran.

E-mail:

gh_medadian@yahoo.com

Dariush Nezhadansari Mahabadi

Assistant professor of language teaching at the University of Isfahan, Iran.

Academic interests

reflective teaching, translation teaching, translation quality assessment, translation contact.

Address

Faculty of Foreign languages, University of Isfahan, Isfahan, Iran.

E-mail:

nejadansari@gmail.com