

# 1 A Comparative Approach to Assessing Assessment

## Revising the Scoring Chart for the Authorized Translator's Examination in Finland

*Leena Salmi and Marja Kivilehto*

### 1.1 Introduction

There is a broad scope of translation quality assessment, and it can occur in various contexts: translator training, machine translation and technical communication, to give a few examples (e.g. Angelelli and Jacobson 2009; Gouadec 2010). Our focus in this chapter is on assessment related to translator certification: the assessment system in the Authorized Translator's Examination in Finland. This examination determines whether the examinees have the professional competence needed for producing so-called "official" or "certified" translations (i.e. legally valid for certain purposes and institutions), but the context can also be considered pedagogical to some extent, since the examinees who fail receive feedback that highlights their abilities and shortcomings (cf. Saldanha and O'Brien 2013, 96).

In our previous articles (Kivilehto and Salmi 2017; Salmi and Kivilehto 2018), we have discussed the assessment system of the Finnish Authorized Translator's Examination and compared it to assessment systems in other certification examinations. The system itself, in its present form since 2008, has been described by Salmi and Penttilä (2013), and is the topic of a recent publication (in Finnish) by the Finnish National Agency for Education (EDUFI) (see Hiirikoski 2017; Kemppanen 2017; Mietunen 2017). We have also examined how a sample of translations have been assessed in the examination. Our purpose has been to gain more information for developing the assessment system of the Finnish examination to make it more valid and reliable, as we have noticed some validity- and reliability-related problems in how the assessment has been applied to translations in the examination (Kivilehto and Salmi 2017).

The scoring chart for assessing translations in the examination was revised in 2017. Previously, the chart comprised two parts, and the translations were marked for both content (C-errors) and language quality (A-errors). The scoring chart currently has three parts, taking into account task accomplishment (T-errors), equivalence of content (C-errors),

and acceptability and readability (A-errors). The previous chart contained 14 error types (see the full chart in Kivilehto and Salmi 2017), while the new one contained only seven.

This chapter presents a comparison of the scoring charts before and after 2017. We describe our process of revising the scoring chart and analyze how it has been applied. The purpose of the comparison is to see if the assessment in the examinations should be further developed, and if so, in what way. In our earlier work, we have discussed similar systems in use elsewhere (e.g. Norway, Germany, Canada or the United States; see Kivilehto and Salmi 2017; Salmi and Kivilehto 2018), and this chapter describes that of the Australian National Accreditation Authority for Translators and Interpreters (NAATI), as a possible way of developing the assessment.

## 1.2 Translation Quality Assessment in Translator's Examinations

### 1.2.1 *Assessment Practices*

Translation quality assessment can be product-, process- and/or user-oriented. Product-oriented assessment is usually based on text analysis and comparing source and target texts (Saldanha and O'Brien 2013, 98–99). One of the best-known, text-based models is that of House (2015), who approaches assessment from the perspective of systemic-functional linguistics and calls her model functional-pragmatic. The principal assessment criterion in House's model is functional equivalence, which can only be reached in translation that is not source culture dependent, i.e. covert translation (House 2015, 60). Otherwise what we are dealing with are different kinds of versions (House 2015, 59). As for process-oriented assessment, it takes a holistic approach to assessment and emphasizes contextual factors such as translator competence and the context in which translations are produced. Examples of process-oriented assessment systems are standards such as the ISO 17100. User-oriented assessment, for its part, focuses on factors such as readability, acceptability and usability, and approaches assessment from an individual's point of view. This means that assessment is related to individual user attributes: reading skills and motivation for reading the translation (Saldanha and O'Brien 2013, 99–100). User-oriented assessment is taken one step further by Suojanen, Koskinen and Tuominen (2015), who introduce practical methods for user-centered translation.

When assessing translations, it is recommended to pay attention to the assessment setting, those doing the translation and the genre and purpose of the translation (House 2015). This applies to examination contexts as well. In the case of the Authorized Translator's Examination in Finland, the texts to be translated fall into the category of legal texts, and thus special attention must be paid to strategies of translating

legal texts. According to Vanden Bulcke and Héroguel (2011, 241), four aspects should be taken into account when assessing translations of legal texts: legal texts as category, genre characteristics, text function and translation strategies. When it comes to certified translations, they often fall into the category of judicial texts (e.g. summons, pronouncements and judgments) and texts that are applications of law (e.g. official documents, contracts and wills). This implies that translations are to be authentic translations that describe the reality of the source text (ST) as closely as possible (Vanden Bulcke and Héroguel 2011, 234, 243). Undoubtedly, the translations must be comprehensible for end users, but as the end users are often experts of the field in question, they may be expected to have the prior knowledge needed for interpreting legal texts of different legal systems. Authenticity, for its part, amounts to foreignizing as a translation strategy. Translations should correspond to STs as closely as possible even with regard to macro- and microstructures, i.e. text structure, phraseology, terminology, syntax and style (Vanden Bulcke and Héroguel 2011, 214). This view is also shared in studies with Danish lawyers and legal translators as informants (Hjort-Pedersen 2016).

Assessment models can roughly be categorized as analytical or holistic (Lommel et al. 2015). *Analytic assessment* focuses “on the identification of precise issues within the object being assessed, such as (for a translation) identification of specific mistranslations, spelling errors,” whereas *holistic assessment* emphasizes “overall characteristics of the object being assessed, such as (in the case of translated texts) reader impression, sentiment, clarity, accuracy, style, whether it enables a task to be completed, and so forth” (Lommel et al. 2015, Section 1.3.2). In assessing certified translation, it is justified to use an analytic rather than a holistic model, since precision is highly valued. Analytic assessment often results in error analysis rather than a comparison of the translation against “ideal” criteria that describe either what the translation should be like or the translation skills it should demonstrate (Angelelli 2009, 40–41; Turner, Lai, and Huang 2010). Error analysis has been regarded as a valid way of measuring translation quality, and this is why it is used in many certification examinations (cf. Hale et al. 2012, 58). An example of this is the certification examination managed by the American Translators Association (ATA 2017). Nevertheless, criterion-referenced assessment can be as valid as error-based assessment (Turner, Lai, and Huang 2010), and at least one certification system, that of the Australian NAATI, has adopted criterion-referenced assessment.

### 1.2.2 Assessment in NAATI Certification Examination

In this section, we discuss the assessment of the Australian certification examination of NAATI. The reason for choosing the NAATI assessment is the fact that it is criterion-referenced, and we see this as one possibility

for developing the assessment of the Finnish Authorized Translator's Examination.

In Australia, the certification examination offered by NAATI takes place several times a year in different language combinations (NAATI 2020a) and has three levels of translator certifications: Certified Advanced Translator, Certified Translator and Recognized Practicing Translator (NAATI 2020b). The Certified Advanced Translator test consists of three tasks: two translations of texts of 400 words and one revision of a translation of 400 words. All STs are written by specialists for specialist readers. They can be research papers, legal briefs or trade agreements, to name a few examples. The test duration is eight hours (NAATI 2020c). The Certified Translator test consists of two translation tasks and one revision task, but the STs are non-specialized texts and shorter (about 250 words) than those in the Certified Advanced Translator test, and they deal with different topics and represent different domains. The domains range from government, legal, health, technology and science to business, society, culture, social services and immigration. The test duration is three and a half hours. In both tests, computers may be used and all kinds of reference materials are allowed. However, neither the use of the Internet nor contacting other people is permitted (NAATI 2020d). For Recognized Practicing Translators, there is no certification test.

The assessment methods of both the translation and revision tasks are criterion-referenced. Two criteria are applied, which means that two competencies are assessed: transfer competency and language competency. For translation tasks, transfer competency means competency in transferring the meaning of the ST, following the translation brief and applying textual norms and conventions, whereas for revision tasks, it means revision skills and competency in applying knowledge of translation standards. As regards language competency, it includes language skills enabling the transfer of meaning. The assessment criteria are considered at five levels, called Bands, of which 1 is the highest and 5 the lowest level. To pass the test, examinees need to achieve at least Band 2 (in some cases 3) for each criterion (NAATI 2020d, 2020e). Bands 2 and 3 for transfer competency and language competency in the translation test for Certified Translator are explained in Table 1.1 (cf. NAATI 2020e).

The overall criteria for NAATI translation examinations are wider than the two criteria described earlier. The NAATI criteria, or competencies, are nine in total, ranging from competencies in transfer, language, research and domain and document types to intercultural, thematic and technological competencies. The examinees are expected to satisfy some prerequisites before taking the certification examination. Prerequisite screening tests are organized in language competency (English proficiency), ethical competency and intercultural competency (NAATI 2020c, 2020e).

Table 1.1 Transfer competency and language competency for Certified Translator in the NAATI translation test (NAATI 2020c)

Transfer competency		Language competency	
Meaning transfer skill	Follow translation brief	Application of textual norms and conventions	Language proficiency enabling meaning transfer
<p>Pass requirements</p> <p>Band 4 </p> <p>At least Band 2</p> <p>Translates the propositional content and intent of the message, with few instances of <b>minor</b> unjustified omissions, insertions and/or distortions. <b>Mostly</b> demonstrates ability to resolve most translation problems appropriately.</p>	<p>At least Band 2 in one of the two criteria, and at least Band 3 in the other</p> <p>Follows the specifications provided in the translation brief. Produces a text which <b>mostly</b> takes into account the purpose of the target text, a specified audience and type of communication.</p>	<p>Demonstrates ability in the use of register, style and text structure appropriate to the genre and <b>mostly</b> consistent with the norms and conventions of the target language.</p>	<p>At least Band 2</p> <p><b>Mostly</b> uses written language competently and idiomatically, in accordance with the norms of the target language. <b>Mostly</b> demonstrates competent use of lexicon, grammar and syntax, including orthography, punctuation and terminology. The target text contains only a few minor errors that <b>do not impact on understanding</b>.</p>
<p>Band 3</p> <p>Translates the propositional content and intent of the message, with several <b>minor</b> and/or any <b>major</b> unjustified omissions, insertions and/or distortions. Demonstrates <b>some</b> ability to resolve translation problems appropriately.</p>	<p>Demonstrates <b>some</b> ability to follow the specifications provided in the translation brief, but does not in several instances take into account the purpose of the target text, a specified audience or type of communication.</p>	<p>Demonstrates <b>some</b> ability in the use of register, style and text structure appropriate to the genre and consistent with the norms and conventions of the target language.</p>	<p>Demonstrates <b>some</b> ability to use written language idiomatically, in accordance with the norms of the target language. Demonstrates <b>some</b> ability to use lexicon, grammar and syntax, including orthography, punctuation and terminology. The target text contains several errors that <b>impact on understanding</b>.</p>

In recruiting assessors, NAATI looks for examiners who have a NAATI certification at an appropriate level and a tertiary qualification in translating, interpreting, language, linguistics or a related discipline. In addition, they have to have near-native competence in the languages they assess, extensive professional experience as a translator or interpreter, commitment to ethical practice and ability to work with others (NAATI 2020f). Each task is assessed by two assessors who work independently. If the assessors disagree with each other about the examinee's performance, i.e. whether s/he should pass or fail, additional assessors will be brought in (NAATI 2020g).

### 1.3 The Finnish Authorized Translator's Examination

#### 1.3.1 Overview of the Examination

In Finland, the EDUFI<sup>1</sup> is the authority that grants the status of an authorized translator with the right to produce legally valid translations, after an applicant has passed the Authorized Translator's Examination or has obtained a Master's degree in translation that includes at least six ECTS credits in certified translation (L 1231/2007). The system is managed centrally for the sake of uniformity, impartiality and equality for the examinees. In the examination, several language combinations are possible, depending on the number of examinees who wish to be tested in a particular language pair and on the availability of qualified assessors. The possibility of becoming authorized on the basis of a university degree in translation, though, applies only to the language combinations available in translator training programs in Finland and only in translation into Finnish or Swedish, including between these two languages. In addition, the status can be granted on the basis of university studies in translation into the student's first language only.

No prerequisites are set concerning the educational background of those wishing to take part in the examination. The examination is offered once a year, usually in November. It tests the examinees' competency in the professional practice of authorized translator and their translation competency in two specialist fields (EDUFI 2019a). In sum, the examination consists of three parts:

- 1 a multiple-choice test on the professional practice of authorized translators (45 minutes);
- 2 one translation assignment in the field of law and administration (2 hours 45 minutes);
- 3 one translation assignment in a field chosen by the examinee (business and economics, medicine, technology or education; 2 hours 45 minutes).

Computers are allowed during the examination. Internet sources and other reference materials may be used during the translation tests. However, the use of translation memories, machine translation and email is not allowed, nor may examinees contact other people (EDUFI 2019a).

### 1.3.2 Assessment in the Examination

In the examination, both language and translation skills are examined using two translation assignments. The assignments are assessed by two assessors. One of them is an expert in the source language (SL) and the other in the target language (TL), though both should be somewhat familiar with both languages. The assessment is performed in accordance with the assessment criteria for language and translation skills (FNBE 2012, 8). The assessors perform the assessments individually, but not completely independently: they are expected to discuss their individual assessments and come to a shared conclusion. If they cannot agree, an additional assessor is usually brought in. To ensure a fair assessment, it is important that the assessment criteria are transparent and consistent, and that the examinees know how the assessment system has been applied to their translations (EDUFI 2019b). The examinees do not have the possibility to appeal to a higher body if they are not satisfied with the assessment, but they have the right to ask the Authorized Translators' Examination Board to reassess their translations (FNBE 2012, 9).

Regarding the qualifications of the assessors, they must have at least a Master's degree and a sound knowledge of translating pragmatic texts in the examination languages. In exceptional cases, a Bachelor's degree may be accepted instead of a Master's degree if the person is a native speaker of the TL. In addition to these criteria, assessors must have completed an assessor training acknowledged by the Finnish National Agency for Education. They are entered into the assessor register for five years; examiners may renew their assessor status, provided that they still satisfy the criteria and have been maintaining their assessment skills (A 1232/2007, Section 12; L 1231/2007, Section 14).

The assessment of translation assignments in the examination is based on an error analysis. The assessors verify how well the source and target texts correspond to each other and how acceptable the translations are as target texts. The first element is generally known as *accuracy* or *adequacy* (Toury 2012, 79), and it relates to what the Multi-Dimensional Quality Metrics (MQM) framework defines as the "extent to which the informational content conveyed by a target text matches that of the source text" (Lommel et al. 2015, Section 1.2). The second element, known in the MQM as *fluency*, refers to "properties of the target text such as grammar, spelling, and cohesion" (Koby et al. 2014, 415).

### 1.3.3 *Revision of the Assessment System*

From 2008 to 2017, the assessors applied a scoring chart containing the two categories of accuracy and fluency: accuracy errors were categorized as errors in the equivalence of content (C-errors) and fluency errors as errors in acceptability and readability (A-errors). Both categories included seven error types. Acting both as researchers and members of the Examination Board, we decided to conduct research on how the scoring chart was applied in practice (Kivilehto 2016, 2017; Kivilehto and Salmi 2017). It became clear that some of the error types were used often, while others were used only rarely (Kivilehto 2016; Kivilehto and Salmi 2017). That is why we decided to propose a new, simplified scoring chart.

As members of the Examination Board, we were able to start the revision work. In the course of it, we applied a user-oriented process, which included a survey and usability testing (see Suojanen et al. 2015). Seminars on the preparation and the assessment of the translations are organized three times a year for the examination assessors within the system (see Salmi and Kinnunen 2015, 235). We started the process in November 2016, in one of these seminars, by surveying the views of the assessors on the assessment criteria (reported in Salmi and Kivilehto 2018, 184–185). We also presented two proposals for simplified assessment criteria, one error-based and the other criterion-based, which were tested in the seminar by the participants. The assessors preferred the error-based criteria, and so, in the following seminar in January 2017, we presented two proposals for error-based criteria. In February–May 2017, the proposal favored by the assessors in the seminar was tested by four experienced assessors. In May 2017, we held a seminar on the scoring chart development with both assessors and translator trainers, where some further adjustments were made. Finally, in November, we had a new scoring chart that was applied in the examination of 2017.

The new scoring chart contains three error categories and seven error types. The errors are categorized as T-errors, C-errors and A-errors. T-errors related to errors in task accomplishment, C-errors related to equivalence of content, and A-errors related to acceptability and readability of text. The category of T-errors was introduced so that special attention could be paid to the characteristics of certified translations such as providing an appropriate heading for the translation and the use of translators' notes. Table 1.2 shows the error types in the new scoring chart.

To compare the previous chart (see Kivilehto and Salmi 2017, Appendix) with the new one, the previous seven A-errors have been combined into two types. As for C-errors, the earlier division into types was more primarily based on where the error occurred: on the level of sentence, grammatical structure or term (e.g. “C2 – A wrong term leading

Table 1.2 The scoring chart in the Authorized Translator's Examination in Finland

<i>T-errors</i> <i>Task accomplishment</i>	<i>C-errors</i> <i>Equivalence of content</i>	<i>A-errors</i> <i>Acceptability and readability of text</i>
T1 The translation has <b>no heading</b> that identifies it as a translation. (5 error points)	C1 <b>Omission</b> (word, term, reference relation or a wider entity). (2, 5 or 10 points)	A1 The <b>syntax, morphology, style, register or idiomaticity</b> of the translation does not follow the norms of the target language. (2, 5 or 10 points)
T2 The <b>function</b> of the translation has been <b>disregarded</b> . (2, 5 or 10 points)	C2 <b>Insertion</b> (word, term, reference relation or a wider entity). (2, 5 or 10 points)	A2 <b>Spelling or orthography</b> does not follow the norms of the target language. (2 points)
	C3 A word, term or a wider entity <b>does not correspond</b> to the source text. (2, 5 or 10 points)	

to the misinterpretation of the translation – 9 error points” or “C6 – Misinterpreted structure – 6 to 2 error points”). The chart also listed separately critical errors that would lead to failing the examination, for example, the C2 error mentioned earlier, or leaving out an entire sentence (C1). The new chart (Table 1.2) focuses more on what the error is like: an element is missing, an element has been inserted or an element does not otherwise correspond to the ST.

In the previous chart, error severity was combined with error type and each error category contained information on how many error points could be given. In the new chart, the severity of errors has three levels: minor (2 points), severe (5 points) and critical (10 points). To fail a translation, only one critical error is enough. The maximum score allowed ranges from 25 to 30, depending on the difficulty of the translation assignment. If an error recurs consistently throughout the text, it is penalized only once.

#### 1.4 Assessment in Practice

In this section, we present an analysis of how the new scoring chart has been applied. The material analyzed contains all the assessed translations from one language setting, the candidates in the language pairs

English-Finnish and Finnish-English, in 2017 and 2018 when the new scoring chart was in application. The data analyzed here consist of translations by nine examinees in the English-Finnish language pair and by 19 in the Finnish-English language pair, altogether 28 examinees. Coincidentally, the number of examinees is the same as in our previous study with data from 2012 to 2014 (Kivilehto and Salmi 2017, 63). Since each examinee produces two translations, there are 56 translations, and as each translation is assessed by two assessors, the data include 112 assessments. Table 1.3 presents the number of translations analyzed.

Three different assessors were involved, all of whom have several years of experience in the examination, and who had also done the assessments analyzed in our previous study (Kivilehto and Salmi 2017). Table 1.4 shows the different error types per language pair, while Table 1.5 presents the results per examination year.

As can be seen in Table 1.4, the most common error type is C3 (a word, term or wider entity that does not correspond to the ST), with 1,112 occurrences. This is similar to our earlier study where the most common type was C7, “an individual word/term that is imprecise, unsuitable or irrelevant or an omission or an addition not essentially affecting the

Table 1.3 Number of translations analyzed, by language pair

	2017			2018		
	<i>Examinees</i>	<i>Translations</i>	<i>Assessments</i>	<i>Examinees</i>	<i>Translations</i>	<i>Assessments</i>
English-Finnish	6	12	24	3	6	12
Finnish-English	8	16	32	11	22	44
Total	14	28	56	14	28	56

Table 1.4 Error types marked in the English-Finnish and Finnish-English translations

<i>Error type</i>	<i>ENG-FIN</i>		<i>FIN-ENG</i>		<i>All</i>	
	<i>Number</i>	<i>%</i>	<i>Number</i>	<i>%</i>	<i>Number</i>	<i>%</i>
C1	28	3.4	69	4.2	97	3.9
C2	8	1.0	46	2.8	54	2.2
C3	474	56.9	638	39.2	1,112	45.2
A1	166	19.9	545	33.5	711	28.9
A2	110	13.2	171	10.5	281	11.4
T1	17	2.0	68	4.2	85	3.5
T2	30	3.6	89	5.5	119	4.8
Total	833	100.0	1,626	100.0	2,459	100.0

meaning of the text” (Kivilehto and Salmi 2017, 63). In both categories, terminology errors explain the high frequency of the most commonly used error type. As we noted earlier, “producing legally valid translations requires accuracy and precision” and the examination texts are “LSP texts that usually contain specific terminology” (2017, 66). Terminology errors are also the explanation given by two experienced assessors (Hiirikoski 2017, 45–46; Miettunen 2017, 72): the translation may be grammatically correct, but the terminology used by the examinee is not the one used in the special field in question, or the examinee fails to recognize the terms used in the ST (see also Kivilehto 2017).

Again, as in our earlier study, the next most common error type is an acceptability error A1 (the syntax, morphology, style, register or idiomat�icity of the translation does not follow the norms of the TL). In the previous scoring chart, this was error A5, “structural error not causing misinterpretation,” which falls within the scope of the current error type A1.

The least often used error type is C2, insertions. Contrary to the earlier study, there are no error types that do not occur at all in this data. This was, in fact, one of the goals of the revision: to have a scoring chart with no error types that are never used. In the data from 2012 to 2014, an error type that did not occur at all was C3, described as “the translation function is disregarded, leading to an inadequate result” (Kivilehto and Salmi 2017, 64). In the present data, insertion (type C2) is the least frequent error type, with 54 occurrences altogether.

Table 1.4 also shows that more errors have been marked in the translations into English ( $n = 1,626$ ) than into Finnish ( $n = 833$ ). This is consistent with our earlier results from 2012 to 2014 (Kivilehto and Salmi 2017, 63). However, as Table 1.3 shows, there were more examinees translating into English (8 in 2017, 11 in 2018) than into Finnish (6 in 2017, 3 in 2018), resulting in 76 assessed translations into English and 36 into Finnish. This amounts to 21.4 errors on average in the translations into English and 23.1 errors into Finnish, so there seems to be no significant difference in the average number of errors.

As regards the distribution of error types per year (Table 1.5), although the number of examinees in both 2017 and 2018 was the same (14), more errors were marked in 2018. There was an increase in all error types except the most common one, C3. The quantitative analysis conducted for this study cannot give a straightforward explanation for this. However, as mentioned, the majority of examinees in 2018 translated from Finnish into English (11 out of 14, see Table 1.3). As Hiirikoski (2017) points out in his analysis of 107 translations into Finnish and 119 translations into English in 2008–2015, translating in this direction seemed to account for more errors than translating into Finnish. We do not have information on the examinees’ linguistic background (mother tongue or other language skills) or their competence in translating that

Table 1.5 Error types marked per year and language direction

2017					2018			
Error type	Number	%	ENG-FIN	FIN-ENG	Number	%	ENG-FIN	FIN-ENG
C1	23	2.1	10	13	74	5.4	18	56
C2	11	1.0	1	10	43	3.2	7	36
C3	565	51.5	371	194	547	40.2	103	444
A1	279	25.4	88	191	432	31.7	78	354
A2	120	10.9	80	40	161	11.8	30	131
T1	27	2.5	12	15	58	4.3	5	53
T2	73	6.6	14	59	46	3.4	16	30
Total	1,098	100.0	576	522	1,361	100.0	257	1,104

might explain the differences, as the examinees are not asked to provide such information. We do not either have exact numbers of how many examinees have passed in each language pair, and statistics published by EDUFI (2019c) show that the overall passing rate is practically the same in both years, 16.9 (12 examinees) in 2017 and 16.7 (11 examinees) in 2018.

### 1.5 Discussion and Conclusions

The comparison of the results on the application of the previous and the new scoring charts shows that the error types most often used are similar: problems with terminology and with the target text syntax (in the earlier results) or its syntax, morphology, style, register or idiomaticity (in the latest results). What is different is that all error types were used with the new scoring chart, contrary to the data from 2012 to 2014. In addition, the assessors' comments on the new chart have been positive. Already during the test phase described in Section 1.3.3, two of the experienced assessors explicitly stated in their written comments that the new system is clear and easy to use, and makes it easier to select the error type. We have not made a comparison of the point scores given, because the scale in the earlier scoring chart was from 1 to 9 as opposed to 2, 5 or 10 points in the new one. However, according to the feedback gathered from the assessors three times a year, there is no reason to believe that the change in the scale has led to failing a translation that would have been accepted using the previous scale.

The aim of our studies in this area has been to increase knowledge of the assessment of translations in examination settings, as well as to contribute to developing the assessment of the Finnish Authorized Translator's Examination. As terminology clearly seems to stand out as a problematic area, separating it as a category of its own might be

considered when developing the error-based scoring chart in the future. In fact, in the MQM typology (Lommel et al. 2015), “terminology” is a category of its own, separate from “accuracy” and “fluency.”

Another idea for developing the examination might be to consider the introduction of two levels of competence, as is the case in the NAATI system described earlier. The statistics (EDUFI 2019c) show that the passing rates of the examination have varied between 8% and 29.9%. The acceptance threshold must be high, as the examination is used to “sort the wheat from the chaff.” As Miettunen (2017, 74) puts it, the resulting translation must be “nearly errorless.” The translations approved need to render the contents of the ST and reflect the legal system of the source culture, and they need to be accurate and precise. Yet, although all translators are not always able to produce a “nearly errorless” translation, they may be able to produce a translation that is suitable for some other purpose than serving as a legally valid document. Therefore, an authorization with two levels of competence might be an idea worth considering: one level for producing legally valid translations that need to reflect the legal system of the source culture, and the other level for translating “in general,” where localizing the text for the target reader is possible. It might also be worth thinking about replacing one of the translation tests with a revision test in line with NAATI (see Section 1.2.2). In real situations where no *one* person has all the knowledge needed, cooperation between a Finnish native translator and an English native reviser (or vice versa) could be a solution. Revision is also a compulsory part of the translation process described in the ISO 17100 standard, and revising skills are equally relevant for the growing practice of post-editing of machine translation. Therefore, it might be a good idea to test revision skills in the examinations.

The assessment of translations in the NAATI examination is criterion-based, which shows that criterion-based assessment can function in certification settings. However, an error-based system also enables detailed feedback to the examinees about the issues that are problematic in their translations. Should a switch to criterion-based assessment be considered, it should be ensured that similar feedback can still be provided to the examinees.

A survey to practicing authorized translators was conducted in 2018 (see Oksanen and Santalahti 2020), and another one was carried out among those who took the exam in November 2019 (see Kivilehto 2020). The translations in the examination are assessed anonymously, and we do not have background information on the examinees (for example, whether they are translating into their first or second language), but we hope to explore their educational background and their reasons for taking the exam. The experienced assessors have the gut feeling that there are candidates who are bilingual or subject experts in a field, but do not have a background in language studies or translating, and who come

to “try their luck” (Kemppanen 2017, 58–59; also Hiirikoski 2017, 43). Kemppanen (2017, 58) also points out that as translators with university training now have the possibility of getting the official accreditation in the language pair of their translation studies, the examinees are more likely not to be formally trained in translation and therefore may not have the necessary skills to pass the exam.

Our plans also include a qualitative analysis on the application of the new scoring chart (Kivilehto 2019), to shed light on the difference shown in Table 1.5. Assessment is always subjective, at least to some extent. However, when assessment criteria and error classification categories are clearly defined, this hopefully leads to more uniformity, impartiality and equality in assessment. The criteria should be comprehensible and easy to use, and we believe that this can best be achieved in cooperation with those who apply them – by taking a user-centered perspective.

## Note

1 Up to 2017, the EDUFI used the English abbreviation FNBE.

## References

- A 1232/2007 = *Valtioneuvoston asetus auktorisoiduista kääntäjistä* [Government Decree on Authorised Translators]. Accessed June 30, 2019. <http://www.finlex.fi/fi/laki/ajantasa/2007/20071232>.
- Angelelli, Claudia V. 2009. “Using a rubric to assess translation ability: Defining the construct.” In *Testing and Assessment in Translation and Interpreting Studies: A Call for Dialogue between Research and Practice*, edited by Claudia V. Angelelli and Holly E. Jacobson, 13–47. Amsterdam: John Benjamins.
- Angelelli, Claudia V., and Holly E. Jacobson, eds. 2009. *Testing and Assessment in Translation and Interpreting Studies: A Call for Dialogue between Research and Practice*. Amsterdam: John Benjamins.
- ATA (American Translators Association). 2017. *ATA Certification Program Framework for Standardized Error Marking Version 2017*. Accessed June 30, 2019. [https://www.atanet.org/certification/Framework\\_2017.pdf](https://www.atanet.org/certification/Framework_2017.pdf).
- EDUFI (The Finnish National Agency for Education). 2019a. *Auktorisoidun kääntäjän tutkinto* [The Authorized Translator’s Examination]. Accessed July 30, 2019. <https://www.oph.fi/fi/palvelut/auktorisoidun-kaantajan-tutkinto>.
- EDUFI (The Finnish National Agency for Education). 2019b. *Arvioijan käsikirja 2018. Auktorisoidun kääntäjän tutkinto. Toimintaohjeet tutkintotehtävien arvioijalle* [Handbook for Assessors 2018. Authorized Translator’s Examination. Instruction for Assessors of the Examination Assignments]. Helsinki: Opetushallitus.
- EDUFI (The Finnish National Agency for Education). 2019c. *Auktorisoidun kääntäjän tutkinnon tulokset 2008–2018* [Results of the Authorized Translator’s Examinations 2008–2018]. Accessed September 29, 2019. PDF available via <https://www.oph.fi/fi/palvelut/auktorisoidun-kaantajan-tutkinto>.

- FNBE (The Finnish National Board of Education). 2012. *Qualification Requirements for Authorized Translators' Examinations 2012*. Regulations and Guidelines 22/011/2012. Accessed June 30, 2019. [https://www.oph.fi/download/191792\\_qualificationrequirements.pdf](https://www.oph.fi/download/191792_qualificationrequirements.pdf).
- Gouadec, Daniel. 2010. "Quality in translation." In *Handbook of Translation Studies*, edited by Yves Gambier, and Luc van Doorslaer, 270–275. Amsterdam: John Benjamins.
- Hale, Sandra B., Ignacio Garcia, Jim Hlavac, Mira Kim, Miranda Lai, Barry Turner, and Helen Slatyer. 2012. *Improvements to NAATI Testing. Development of a Conceptual Overview for a New Model for NAATI Standards, Testing and Assessment. The National Accreditation Authority for Translators and Interpreters (NAATI)*. Accessed March 13, 2020. <https://www.naati.com.au/media/1062/intfinalreport.pdf>.
- Hiirikoski, Juhani. 2017. "Lain ja hallinnon käännoستهتävät englannin kielessä" [Translation tasks in the special field law and administration, with reference to translating from and into English.] In *Auktorisoidun kääntäjän tutkinnon historiaa ja nykypäivää* [The Authorized Translator's Examination: Past and Present], edited by Tarja Leblay, 37–48. Opetushallitus, Raportit ja selvitykset 2017:16. Accessed March 13, 2020. <https://www.oph.fi/fi/tilastot-ja-julkaisut/julkaisut/auktorisoidun-kaantajan-tutkinnon-historiaa-ja-nykypaivaa>.
- Hjort-Pedersen, Mette. 2016. "Free vs. faithful – Towards identifying the relationship between academic and professional criteria for legal translation." *English Language Overseas Perspectives and Enquiries* 13(2): 225–239. doi:10.4312/elope.13.2.225-239.
- House, Juliane. 2015. *Translation Quality Assessment. Past and Present*. Abingdon: Routledge.
- Kemppanen, Hannu. 2017. "Auktorisoidun kääntäjän tutkinnon tehtävien laadinta ja arviointi: venäjän kielen näkökulma." [Preparing and assessing translations for the Authorized Translator's Examination: The Russian language perspective.] In *Auktorisoidun kääntäjän tutkinnon historiaa ja nykypäivää* [The Authorized Translator's Examination: Past and Present], edited by Tarja Leblay, 49–60. Opetushallitus, Raportit ja selvitykset 2017:16. Accessed September 29, 2019. <https://www.oph.fi/fi/tilastot-ja-julkaisut/julkaisut/auktorisoidun-kaantajan-tutkinnon-historiaa-ja-nykypaivaa>.
- Kivilehto, Marja. 2016. "Käännösfunktion huomiotta jättäminen, joka johtaa epätäsmälliseen lopputulokseen. Auktorisoidun kääntäjän tutkinnon käännöstehtävien arvioinnista." [The translation function is disregarded, leading to an inadequate result. On assessing translation assignments in the authorised translator's examination]. In *Text and Textuality. VAKKI Publications* 7, edited by Pia Hirvonen, Daniel Rellstab, and Nestori Siponkoski, 391–401. Vaasa: University of Vaasa. Accessed October 4, 2019. [http://www.vakki.net/publications/no7\\_eng.html](http://www.vakki.net/publications/no7_eng.html).
- Kivilehto, Marja. 2017. "Miten auktorisoidun kääntäjän tutkinnon käännöstehtävät vastaavat tutkinnon tavoitteita erikoisalojen kääntämisen näkökulmasta?" [How do the translation assignments in the authorised translator's examination meet the requirements of the examination from the point of view of specialised translation?]. In *MikaEL. Electronic Journal of the KäTu Symposium on Translation and Interpreting Studies*, edited by Ritva Hartama-Heinonen, Marja Kivilehto, Liisa Laukkanen, and Minna Ruokonen, Vol. 10,

- 136–149. Accessed October 4, 2019. <https://www.skyl.fi/liitto/seminaarit/mikael-verkkojulkaisu/>.
- Kivilehto, Marja. 2019. “Mellan uppgifts-och examenskontext. Var befinner sig examinandan?” [Between the assignment and examination contexts. Where is the examinee?]. Presentation at the Conference of *Svenskans beskrivning* 37 in Turku, May 10, 2019.
- Kivilehto, Marja. 2020. “‘Vahvistan, että tämä käänös on ...’. Autenttisuuden dilemma(ko?) auktorisoidun kääntäjän tutkinnossa” [‘I confirm that this translation is ...’. Dilemma (?) of authenticity in Authorized Translator’s Examination]. Presentation at *Symposium 2020 – Workplace Communication III. XXXX International VAKKI Symposium*. February 6–7, 2020 in Vaasa, Finland.
- Kivilehto, Marja, and Leena Salmi. 2017. “Assessing Assessment: The Authorized Translator’s Examination in Finland.” *Linguistica Antverpensis, New Series: Themes in Translation Studies* 16: 57–70.
- Koby, Geoffrey S., Paul Fields, Daryl Hague, Arle Lommel, and Alan Melby. 2014. “Defining translation quality.” *Revista Tradumatica* 12, 413–420.
- L 1231/2007 = *Laki auktorisoiduista kääntäjistä* [Act on authorised translators]. Accessed June 30, 2019. <http://www.finlex.fi/fi/laki/ajantasa/2007/20071231>.
- Lommel, Arle, Aljoscha Burchardt, Attila Görög, Hans Uszkoreit, and Alan Melby, eds. 2015. “Multidimensional quality metrics (MQM) issue types.” Accessed September 28, 2019. <http://www.qt21.eu/mqm-definition/definition-2015-12-30.html>.
- Miettunen, Markku. 2017. “Englanti: kokemuksia talouselämän erikoisalasta.” [English: Experiences on assessing translations in the special field of economics]. In *Auktorisoidun kääntäjän tutkinnon historiaa ja nykypäivää* [The Authorized Translator’s Examination: Past and Present], edited by Tarja Leblay, 67–74. Opetushallitus, Raportit ja selvitykset 2017:16. Accessed September 29, 2019. <https://www.oph.fi/fi/tilastot-ja-julkaisut/julkaisut/auktorisoidun-kaantajan-tutkinnon-historiaa-ja-nykypaivaa>.
- NAATI (National Accreditation Authority for Translators and Interpreters). 2020a. *Certified Translator*. Accessed July 1, 2019. <https://www.naati.com.au/certification/certification-testing/certified-translator/>.
- NAATI (National Accreditation Authority for Translators and Interpreters). 2020b. *Descriptors for Translator Certifications*. Accessed 3 June 30, 2019. <https://www.naati.com.au/media/1586/descriptors-for-translator-certifications-version-1-june-2017.pdf>.
- NAATI (National Accreditation Authority for Translators and Interpreters). 2020c. *Certification Scheme Design Summary*. Accessed June 30, 2019. [https://www.naati.com.au/media/2397/certification-scheme-design-summary\\_may2019.pdf](https://www.naati.com.au/media/2397/certification-scheme-design-summary_may2019.pdf).
- NAATI (National Accreditation Authority for Translators and Interpreters). 2020d. *Certified Translator Test. Candidate Information*. Accessed June 30, 2019. [https://www.naati.com.au/media/2232/ct\\_candidate\\_information.pdf](https://www.naati.com.au/media/2232/ct_candidate_information.pdf).
- NAATI (National Accreditation Authority for Translators and Interpreters). 2020e. *Certified Translator Test Assessment Rubrics*. Accessed June 30, 2019. [https://www.naati.com.au/media/2231/ct\\_assessment\\_rubrics.pdf](https://www.naati.com.au/media/2231/ct_assessment_rubrics.pdf).

- NAATI (National Accreditation Authority for Translators and Interpreters). 2020f. *Expression of Interest. NAATI Examiner Panels*. Accessed June 30, 2019. <https://www.naati.com.au/media/1983/examiner-eoi-info-handout-18pdf.pdf>.
- NAATI (National Accreditation Authority for Translators and Interpreters). 2020g. *How Are Certification Tests Marked?* Accessed June 30, 2019. <https://www.naati.com.au/certification/certification-testing/>.
- Oksanen, Henrik, and Miia Santalahti. 2020. "Auktorisoidun kääntämisen tila 2019. Kyselytutkimus auktorisoidujen käännotösten tekstilajeista ja auktorisoidun kääntäjän ohjeiden käytöstä." [The status of authorized translating in 2019. Survey on text genres of legally valid translations and on using the instructions for authorized translators.] In *MikaEL, Electronic Journal of the KäTu Symposium on Translation and Interpreting Studies*, edited by Ritva Hartama-Heinonen, Laura Ivaska, Marja Kivilehto, and Minna Kujamäki, Vol. 13, 25–42. Accessed May 17, 2020. <https://www.sktl.fi/liitto/seminaarit/mikael-verkkojulkaisu/>.
- Saldanha, Gabriela, and Sharon O'Brien. 2013. *Research Methodologies in Translation Studies*. London & New York: Routledge.
- Salmi, Leena, and Tuija Kinnunen. 2015. "Training translators for accreditation in Finland." *The Interpreter and Translator Trainer*, 9 (2): 229–242.
- Salmi, Leena, and Marja Kivilehto. 2018. "Translation quality assessment: Proposals for developing the authorised translator's examination in Finland." In *Legal Translation and Court Interpreting: Ethical Values, Quality, Competence Training*, edited by Annikki Liimatainen, Arja Nurmi, Marja Kivilehto, Leena Salmi, Anu Viljanmaa, and Melissa Wallace, 179–198. Berlin: Frank & Timme.
- Salmi, Leena, and Ari Penttilä. 2013. "The system of authorizing translators in Finland." In *Assessment Issues in Language Translation and Interpreting*, edited by Dina Tsagari and Roeland van Deemter, 115–130. Frankfurt am Main: Peter Lang.
- Suojanen, Tytti, Kaisa Koskinen, and Tiina Tuominen. 2015. *User-Centered Translation*. New York: Routledge.
- Toury, Gideon. 2012. *Descriptive Translation Studies – and Beyond*. Revised 2nd edition. Amsterdam/Philadelphia: John Benjamins.
- Turner, Barry, Miranda Lai, and Neng Huang. 2010. "Error Deduction and Descriptors – A Comparison of Two Methods of Translation Test Assessment." *Translation & Interpreting. The International Journal for Translation & Interpreting Research* 2(1): 11–23.
- Vanden Bulcke, Patricia, and Armand Héroguel. 2011. "Quality issues in the field of legal translation." In *Perspectives on Translation Quality*, edited by Ilse Depraetere, 211–248. Berlin & Boston: De Gruyter Mouton.