Usability Methods in Translation Evaluation: Heuristic Evaluation and Usability Testing

Juho Suokas, University of Eastern Finland

ABSTRACT

This article examines the application of usability methods in translation evaluation. The methods are used to analyse a translation criticised for its difficult language and poor translation quality. Instead of traditional translation quality assessment, the focus is on examining the translation's usability. Usability evaluation places the focus on the user and use-context of the product. In this study the Finnish translation of The Guitar Handbook is evaluated by using heuristic evaluation and usability testing. Both methods would seem to be well suited to translation evaluation, pointing out important usability issues in the text and providing two different points of view. Usability evaluation would appear to offer new and interesting angles for evaluating translations and developing translation quality assessment models.

KEY WORDS: usability, user-centred translation, heuristic evaluation, usability testing, translation quality assessment

1 INTRODUCTION

Translation evaluation is not always a straightforward task. Translation quality is a constant subject of debate in both translation practice and translation studies (TS). There are various conflicting opinions on what exactly quality is and how to define it. Nonetheless, different texts have various uses and diverse groups of readers, who in turn have different needs when it comes to the text itself. If we were to evaluate translations and how well they meet the needs of their reader, could there be another way of evaluating them instead of focusing on traditional translation quality assessment? This paper

presents a small-scale study using usability methods as an alternative means of evaluation.

The text material chosen for this study is an instructional book, which has been criticised for containing poorly translated language. When discussing poor quality translations in the Finnish newspaper *Helsingin Sanomat* (weekly supplement *Nyt* 2000:61), Tero Valkonen, a professional translator himself, uses Ralph Denyer's *The Guitar Handbook* (1982) as a prime example of how not to translate into Finnish. He harshly criticises the translation's language and describes how 'the Finnish reader must fight their way through impossible language to get to the point' (my translation). This quote could be seen to represent a usability issue with the translation – the reader is using the book to achieve some specific goal and the language makes achieving it problematic, thus hindering the book's usability.

By taking advantage of methods from usability research, Valkonen's comment can be used as a starting point for a small-scale study of translation usability. As a language professional, Valkonen is well acquainted with finding problems in translations and their language quality. However, if we were to focus on usability and actual Finnish readers of the translation, would similar problems still surface? Here I present a case study using heuristic evaluation and usability testing to evaluate *Suuri kitarakirja*, the Finnish translation of *The Guitar Handbook*. This article is divided into seven sections, starting with this introduction. The second section addresses the concept of usability on a general level. The third section looks at methods of usability evaluation. The fourth section describes how usability methods have been applied in two experimental tests used to evaluate an excerpt from *Suuri kitarakirja*. The fifth section presents the results of these two tests. The results are then discussed in the sixth section, and the final section offers a general conclusion.

2 USABILITY

Jakob Nielsen (1993:25) — one of the best-known names in usability research — places usability under the broader concept of usefulness, where a product's usefulness consists of two parts: utility, i.e. the product can be used to achieve a certain goal, and usability, the ease of which these goal(s) can be achieved. There are many definitions of usability. Notably, ISO 9241-11 defines usability as '[t]he extent to which a product can be used by specified users to achieve specified goals with effectiveness, efficiency and satisfaction in a specified context of use.' (ISO 1998.) From the definition we can see that the focus is on the situation of use, and that the users and goals need to be specified. While the effectiveness and efficiency are factors that can be measured externally by studying a use situation, it is noteworthy to mention that user satisfaction can only be analysed by learning what the users themselves think. Nielsen (1993:23) also adds that when the term usability is used, the methods for improving the product are often included in the term.

The concept of usability originates from Human–Computer Interaction studies, where it has been used as a quality component to evaluate and develop user interfaces (Suojanen et al. 2015:14). However, since then the study of usability has expanded to cover other types of interfaces as well, including texts. We can see usability as a feature of any item

designed for a specific purpose, anything from can openers to spacecraft. Imagine a square-shaped wheel and you have a product with a clear usability problem. Various terms have been used to discuss similar topics in the user experience (UX) and user-friendliness, such as human factors and ergonomics (Nielsen 1993:23).

Translation usability is a fairly new object of study that has not been researched extensively. Some early instances of usability in TS include Leena Salmi's dissertation *Documents multilingues pour logiciels et utilisabilité* (2003) and Jody Byrne's *Technical Translation: Usability Strategies for Translating Technical Documents* (2006). In his book, Byrne suggests that software user guides can be seen as a type of user interface, thus a focus on usability could result in better user guides. Byrne himself defines the specific usability of texts as follows: "When applied to texts usability measures the extent to which readers can read a text, understand its content and perform whatever task is required by the text quickly and accurately and the extent to which they find the experience difficult or easy" (2012:201). Usability is also mentioned by Pym (2011:416-418) as an important concept of website localisation, and it has been used in recent studies to examine machine translation (see e.g. Doherty & O'Brien 2014).

Recently, there has been a growing interest in usability in translation with the introduction of user-centred translation (UCT). The UCT model by Tytti Suojanen, Kaisa Koskinen and Tiina Tuominen (2012; 2015) is based on the fundamentals of usability research. It aims to offer concrete methods to take into account the reader when producing translations, bringing Byrne's ideas about translation and usability to a more general level that is not limited simply to technical translation. The UCT model is designed to be suitable for translation practice as well as translator training and TS in general. UCT has adapted its name and main principles from user-centred design, a model where a product is developed iteratively to meet the needs of its users. UCT presents methods used in usability research that can be applied to translation in order to concretely address the reader. UCT can be seen as a practical evolution of functionalist translation theories, with a focus on the reader/user of the translation (e.g. Suojanen et al. 2012:9).

3 METHODS IN USABILITY RESEARCH

Creating products with high usability requires that the user is taken into account at every stage of the design process. The product is developed based on information gathered during the different stages. This iterative approach is known as user-centred design. There are several practical ways of taking the user into consideration throughout the product's design and development process. These methods are included at various stages of the process, covering each aspect of the development from initial designing to the project's post-mortem evaluation.

Usability evaluation methods can be broadly defined as falling into two categories: empirical methods and expert analysis (Byrne 2006:180; Suojanen et al. 2015:93). Empirical usability methods involve observing actual users of the evaluated product. The most common empirical method in usability research is usability testing, which according to Nielsen (2012:29) and Byrne (2006:180) produces the most relevant information. Usability can also be evaluated without performing empirical tests with actual end users by using

heuristic evaluation instead, which is performed by experts instead of users. These two methods will be presented here in more detail.

3.1 Heuristic evaluation

Heuristic evaluation is often one of the first usability methods used during the early stages of product development, for it is cost-effective and a relatively simple method to apply (Nielsen 1993). The evaluation is based on a set of usability principles, or heuristics. These principles can be presented as a checklist for the evaluators. To give an example of what usability heuristics might look like, here are a few examples of Nielsen's heuristics used in computer interface design:

> Speak the users' language: The dialogue should be expressed clearly in words, phrases, and concepts familiar to the user, rather than in system-oriented terms. […] Minimize the users' memory load: The user should not have to remember information from one part of the dialogue to another. Instructions for use of the system should be visible or easily retrievable whenever appropriate. (Nielsen 1993:20)

The evaluation is carried out by a small group of experts with either knowledge of the product or usability – or both. For this reason, heuristic evaluation is often also called expert evaluation. The evaluators go through the product and look for usability problems, which they list according to the heuristics they violate. Often the problems are also given a severity rating of 0–4 as well as suggestions for improvement. After individual evaluations, the evaluators gather together to discuss their findings and create the final evaluation report. The results of the evaluation are used to improve usability during the product's development process; commonly a heuristic evaluation is performed and the issues found during the evaluation are addressed before moving on to empirical usability tests (Suojanen et al. 2015:78). To give an example of what a heuristic evaluation can look like, a typical usability problem based on the aforementioned Nielsen's heuristics is reported in Table 1 (fictional example):

Table 1. Example of heuristic evaluation report

ID	Location	Problem	Heuristic	Severity	Suggestion
JS_010	Page 2, line 34	Unfamiliar term 'interlingual communication expert'	Speak the user's language	2	Use more common term: 'translator'

There is a clear correlation between heuristics and translation quality principles and style guides used in the translation industry. Indeed, revision is where heuristics could be used beneficially, when applied to translation practice, since many companies already have models of revision with similarities to heuristic evaluation (Suojanen et al. 2012: 109; Lång, 2013).

Heuristic lists can be general or targeted for a specific product. While they are not common as such in translation, some heuristics exist for technical writing, such as Purho's 10 heuristics for documentation (2000). Recently, Suojanen et al. (2015:90) have suggested a list of general UCT heuristics for translation. The UCT heuristics are based on Purho's heuristics, Daniel Gouadec's translation quality principles (2007:6-8), and Nielsen's heuristics (1993:20), as well as on several previous studies in translation usability. However, at the time of study in 2014, general UCT heuristics did not exist, so a new set of translation heuristics was created. These heuristics are presented in section 4.2.

While heuristic evaluation can be cost-effective and address various usability issues (including expert-level usability problems) in the product, the downside is that the method does not involve actual users. Testing with actual users can present unexpected usability problems, which might not be visible from an expert's perspective. Nielsen suggests that using a combination of heuristic evaluation and usability testing is often the most useful solution (1993:223-226). The two methods would appear to complement each other quite well and most of the problems in any single method can be addressed by using the other.

3.2 Usability testing

Usability testing is one of the core methods in usability research. It involves actual test users who perform predetermined tasks with the evaluated product. The main aim is to gather information on how the users actually behave when they utilise the evaluated product. A group of test users are gathered for the evaluation. The participants either belong to the target group of the product or would be likely to use the product. Often three to five test users are enough for each test, for research suggests that even such a small group of participants can find around 80 percent of a product's usability problems (Nielsen 2000; Rubin & Chisnell 2008:72; Suojanen et al. 2015:95).

In a usability test, each user individually performs one or more predetermined tasks with the product while being observed by the test moderator(s). The observation can be direct, i.e. the observer is present while the user performs the task, or indirect by the use of recording the user's actions. Often a combination of both is used. When using only direct observation, the data gathered is completely reliant on the observer and their attention. Recorded material can include video recordings, software logging or eye-tracking (Byrne 2006:182-184; Suojanen et al. 2012:75-76). The users can also be asked to vocalise their thoughts while performing the tasks. This thinking aloud is a common feature of usability tests and it is also used in translation process research (Suojanen et al. 2015:98-100). The observer should not interact with the test user during the test to achieve a realistic outcome of an actual use situation. The observation is often complemented with interviews and questionnaires, either before or after the test session. However, it must be noted that a usability test in itself is an artificially created situation, which might affect the user's performance. In usability research, ethnographic methods can be used to address this issue.

4 THE CASE: METHODS AND MATERIAL

In this study, heuristic evaluation and usability testing have been chosen to evaluate the usability of *Suuri kitarakirja*. As described in the previous chapter, both methods are common ways of assessing usability in usability research and they are also suggested as a useful and cost-effective combination by Nielsen (1993:225-226).

4.1 Material

The Guitar Handbook by Ralph Denyer (1982) is a popular instructional book for guitar players of different levels of expertise. It covers various aspects of guitars and guitar playing, such as notable guitarists, different guitar types, instructions for playing and guides for guitar set up and maintenance. The Finnish translation, *Suuri kitarakirja* (translated by Saastamoinen et al.) on the other hand, has received criticism for the quality of its language, as outlined in the critique by Valkonen presented in the introduction. The source text and target text share the same layout, including page numbers, pictures and typography.

Both heuristic evaluation and usability testing are applied to the same excerpt chosen from *Suuri kitarakirja*. The evaluated section consists of four pages describing basic solo playing techniques (pp. 140-143). The section begins with a one-page overview of improvisation, which includes quotes from famous guitarists, such as Jeff Beck and Ritchie Blackmore. The instructional part of the section describes how to perform four different playing techniques: slurs, sliding, vibrato and bending. Most of the relevant terms are presented in both Finnish and English. The section includes instructional photographs and pictures alongside the text. The passage does not include specific examples presented in Valkonen's critique, but it can be seen to represent the overall style of the book. The excerpt was chosen for its informative description of relatively basic-level playing techniques, which can be well adapted for usability testing.

4.2 Heuristic evaluation

For the purpose of this study, a specific list of heuristics was created. These heuristics (below) are based on Purho's documentation heuristics and Daniel Gouadec's translation quality principles presented in *Translation as a Profession* (2007:6-8).

- Matching the real world
    - The text is compliant with the translation's physical limitations and the target community's rules, regulations as well as linguistic and cultural standards.
- Accessibility
    - The text is well written; its overall language is familiar to the user.
    - The words, phrases and concepts used in the material are familiar and they are used consistently.
- Accuracy
    - The text is as factual and as free of technical and semantic errors as possible.

- Purposeful and ergonomic
    - The function/use of the text is clear to the user.
    - The information is focused on the task at hand.
    - An appropriate medium is used.
- User support
    - The information is suitable for users with different levels of experience.
    - The text provides support for possible problems that might arise while using it.
- Information design
    - The information is easily found and understood.
    - It is presented in a logical and natural way.
    - The paragraph sizes and use of graphics are effective.

The evaluation performed here is a slightly modified version of an average heuristic evaluation used in usability research. The three evaluators were members of the academic staff of the Philosophical Faculty of the University of Eastern Finland, chosen on the basis of their expertise in language and experience with guitar playing. Their academic backgrounds included linguistics, literature and translation. All evaluators stated they had over 20 years of guitar playing experience with varying levels of playing activity. The evaluators also stated they had at least some previous experience with *The Guitar Handbook* or similar learning material. All of the evaluators were native Finnish speakers. They represent expert evaluators who have knowledge of the product under examination, but none of them had significant knowledge of usability research. They will be referred to as evaluators A, B and C.

Since the evaluators had no previous experience in usability evaluation, the list of heuristics was formulated into a set of questions that the evaluators answered based on the text. One further general question was also added to the list, which did not relate directly to heuristics, but asked if any problems were not addressed by the previous questions. The evaluation was carried out in Finnish. In contrast to traditional heuristic evaluation, no severity ratings were given to the usability problems identified and the evaluators did not come together to discuss their findings due to time constraints. The questions are presented here (translated from Finnish) with their corresponding heuristics.

- Are the words, phrases and concepts familiar to the user? Are they used consistently? (Accessibility)
- Does the text deviate from common usage in the target language or community? (Matching the real world)
- Are there factual or linguistic errors? (Accuracy)
- Is the intended function clear to the reader? Does the given information focus on the purpose of use? (Purposeful and ergonomic)
- Is the information suitable for users (players) with diverse levels of experience? (User support)

- Does the text provide support for possible problems that could come up during use? (User support)
- Is the information presented in a logical and natural way? (Information design)
- Is a printed book an appropriate medium for this purpose? (Purposeful and ergonomic)
- Is the structure efficient (layout, use of graphics, paragraph sizes…)? (Information design)
- Does the book seem appealing to use?
- Is the text well written? (Accessibility)
- Did some other specific problems arise concerning the use of the book? (General)

4.3 Usability testing

The test group for the usability testing comprised four students in their 20s from the University of Eastern Finland in Joensuu. In contrast to the expert evaluation, language students and language experts were not chosen for the test. The participants' guitar playing experience levels varied, though all test users stated they had started playing at least eight years previously with various levels of activity. Three of the participants reported they were completely self-taught players, while one had studied guitar briefly at a music institute, but considered himself to be mainly self-taught. Only one of the participants had some previous experience with *Suuri kitarakirja*, but all had used similar self-learning material. Three of the participants were male and one female, but 'he' will be used here as a generic pronoun for all the participants for the purposes of anonymity. They will be referred to as participants 1, 2, 3 and 4.

The usability tests in this study included direct observation and survey methods. The tests were carried out in March 2014 at a rehearsal room in Joensuu. The participants were asked to fill in a background information form and an informed consent form. The participants were given a chair and an electric guitar plugged into an amplifier, the book was placed on a flat surface in front of them. They were asked to go through the text and practise the techniques described therein. The participants were given up to 30 minutes to finish the task while being observed by the test moderator. In this case, the observation was carried out directly without recording the users. While not having video recordings of the test situations does affect the material gathered, this decision was made to make the test situation more comfortable for the test users, so that they would not feel that their guitar playing skills were under evaluation.

A successful completion of the test task was defined as 'the participant understands the text and is able to learn the given techniques from it', or if the participant was already familiar with the techniques, 'the participant understands the text and finds it adequate for practising the given techniques'. All participants were familiar with the techniques, so the latter definition was used. The tests were followed by a semi-structured interview, which was recorded with an audio recorder.

As mentioned previously, usability testing is often conducted after the problems found in the heuristic evaluation have been addressed. This study applied a similar way of beginning the evaluation with heuristics followed by usability testing, but since both methods were used on a finished product, the issues found in heuristic evaluation could

not be addressed before the usability testing. Thus the iterative in-process application of usability methods as suggested by the UCT model does not apply here.

5 RESULTS

5.1 Results of heuristic evaluation

All reported comments by the evaluators and test users are the author's translations from Finnish. First addressed are problems found relating to the first two heuristics: 'Matching the real world', which requires compliance with the translation's physical limitations and the target community's rules, regulations and linguistic and cultural standards, and 'accessibility' which requires that the text should be well written and its overall language should be familiar to the reader, including words, phrases and concepts, which should also be used consistently. These are arguably the most relevant heuristics to consider when addressing the critique of the translation's language quality.

When answering questions relating to the 'matching the real world' heuristic, the evaluators found some deviation from common linguistic features of the Finnish language. According to the evaluators, these included some archaic or awkward features in word choices and syntax. Evaluator C commented that the text gives the impression of being a word-for-word translation and that the overall language is not very good Finnish. Some parts of the text were seen to include source language interference, such as referring to strings with a lighter touch as *kevytkosketuksisempiin kieliin*, which was not seen as an idiomatic Finnish expression. The evaluators also found some unnecessarily noun-heavy expressions such as *kielen venyttäminen suoritetaan* and *tapa välttää kielen epävireeseen menoa*, where the use of the word *kieli* (string) was seen as unnecessary and adding extra weight to the expression.[9]

Regarding the 'accessibility' heuristic, the evaluators did not find the text to be very well written in terms of good Finnish language. However, the evaluators suggested that most terms are clear and familiar to users with some level of musical experience, although modern readers might be more familiar with the original English terms than their translations. A few of the translated terms were considered to be unfamiliar or even amusing; these included such translations as *nauhaisku* and *nyhtäisy*, which in the text are presented alongside their English terms 'hammer on' and 'pull off'. The terminology was seen to be mostly consistent, except for the term 'bending' which was translated as *venyttäminen* but also on occasions as *taivuttaminen*, which one evaluator found to be an uncommon way of referring to the technique.

From a broader perspective, however, the evaluators did not consider the problems with the language features to be too severe. Two of the evaluators stated that the translated version would need work on its language, but the language problems do not affect understanding of the topic. As evaluator A commented, "the text manages to present the information, but it is in no way a great reading experience." One evaluator had taken a look at the original English version of the book and described its language as fluent and

[9] My translations: 'the string bending is performed' and 'a way to avoid the string going out of tune.'

natural. Evaluator C described the Finnish text's tone as "motivating and inspiring, despite its awkwardness."

Examining the 'accuracy' heuristic also highlighted some problems. This heuristic requires the text to be as factual and as free of technical and semantic errors as possible. While the evaluators agreed that the text is factually accurate, they found problems in spelling and grammar, such as missing or misplaced commas and uncommon sentence structures. A few typos were also found.

The 'purposeful and ergonomic' heuristic requires that the function/use of the text must be clear to the reader, the information is focused on the task at hand, and an appropriate medium is used. All evaluators agreed that the function of the text is clear and the information is mostly effectively focused and relevant. However, evaluator B commented that to be more in line with Finnish cultural standards for informative texts, the information should be more direct and to-the-point. Evaluator C found unnecessary information on page 141 (performing a pull-off at a certain angle), which he thought could be more distracting than helpful. Regarding the medium, all evaluators thought that the book format was suitable for the purpose and a good interface, but some audio-visual material would have been helpful to complement the text, since as evaluator B put it, "after all, music is primarily sound."

The next heuristic, 'user support', requires that the text should be suitable for users of different levels of experience and that the text should provide solutions for possible problems that might arise during use. The evaluators thought that while the book as a whole would suit players with various levels of experience, the excerpt used here would be mainly suitable for beginners. The evaluators also noted that some attention had been paid to take account of possible problems during use, such as how to avoid unwanted noise when bending strings.

The final heuristic, 'information design', states that information should be easily found and understood, presented in a logical and natural way and that the paragraph sizes and use of graphics should be effective. The evaluators found the flow of information to be in line with the heuristic, but they criticised the layout, where the information is presented in different sections inside separate boxes. Evaluator A found that this creates a fragmented impression. Evaluator C suggested that the separate sections lose their effectiveness, since the whole body of the text is presented inside these boxes. He also added that the layout and design is not especially vibrant or interesting compared to modern learning material. Similarly, the evaluators found problems with the pictures appearing unclear: the relation between the picture and the text was not seen to be always clearly apparent. All in all, the heuristic evaluation would suggest that the text is understandable, but has linguistic problems as well as a problematic layout.

5.2 Results of the usability test

The usability test pointed out similar problems as the heuristic evaluation, but not all results were alike. This supports the idea of using both heuristic evaluation and usability testing to complement each other, as mentioned in section 3.2.

While performing the task, there were significant differences between how the users used the book to practise the techniques. Two users read through the text one section at a

time and practised the techniques as they came up in the text. One user read through the whole text first before applying the playing techniques and one user skimmed through the text as a whole and then proceeded to practise the techniques in a non-linear order, different to the order they were presented in the text. Similarly, two participants performed the techniques only in the way they were presented in the text, while two participants used the techniques as a part of improvised playing. One participant was improvising on the guitar most of the time while reading the text, others paused reading to play and vice versa. The time the participants spent on the task varied between 10 and 25 minutes. The book seemed quite easy to use placed on a surface in front of the test participants. However, the small font did seem to present some problems, since at times participants had to lean close to the book with the guitar on their lap.

The participants were encouraged to speak aloud if something in the text caught their attention, but they remained mostly quiet during the test. The think aloud method was not used for this test. The few instances when the participants did speak aloud dealt mainly with parts of the text that presented unfamiliar information or parts that were considered amusing – such as the translated term *nauhanyhtäisy*, which participant 3 said he had never encountered before. Participant 4 spoke aloud that he had never performed the hammer-on and pull-off techniques in the way they were presented in the text (as a four note chromatic progression) and thought that it was not very useful for his own style of playing. While reading, participant 4 also criticised the large amount of text being packed into a small space.

While performing the task, some of the participants had trouble performing the vibrato technique horizontally – along the neck – as opposed to vertically, where the string is bent sideways. None of the participants had trouble with the vertical vibrato. Participant Three stated that although the text describes the horizontal technique as the more common one, he had never performed it that way before. He thought it might suit an acoustic classical guitar better than an electric guitar, which was used in the test.

All test participants had successfully completed the test according to the definition *the participant understands the text and finds it adequate for practising the given techniques*. Each participant found the text to be understandable and useful for practising the techniques. Participant 1 commented that the text had given him a chance to brush up on playing techniques that he had not been practising for a while. Participant 2 stated that the text had given him a chance to fine tune the techniques. Participants 3 and 4 stated that while they had not learned anything new from the text, they found it to be suitable for learning and practising the techniques. Still, the interviews did highlight usability problems that the participants had encountered during the test.

In contrast to the first part of the heuristic evaluation where the heuristics 'matching the real world' and 'accessibility' were addressed, none of the participants had negative comments about the language of the text. When asked about the language, all participants found the text to be well written and good Finnish. Participants 1 and 4 had liked the tone of the text and described it as helpful and motivating. Participants 1 and 3 found the educational style of the text to be well-suited for the purpose. The terminology was found to be overall familiar to the participants, especially the English terms which were presented alongside the Finnish ones. However, the Finnish term *nauhanyhtäisy* for 'pull-

off' was new for all participants except participant 2. The others mostly found it amusing. This term had not been used in the music institute in which one of the participants had studied, where only the English terms 'hammer on' and 'pull-off' were used.

Whereas the expert evaluators had found the text to be factually accurate, but containing semantic errors (the 'accuracy' heuristic), the only factual problem the participants found was related to the horizontal vibrato technique. As mentioned above, the horizontal vibrato was described as the more common way of performing the technique, but the participants were more familiar with performing the technique vertically. No other factual or semantic errors were pointed out in the usability testing.

All the participants had indeed found the text to be adequate for practising the given techniques and the function of the text was clear to them, as with the evaluation of the 'purposeful and ergonomic' heuristic. However, the focus of the information presented some problems, concerning mainly the sheer amount of information. Participants 1 and 4 commented that there might be too much content to keep the reader interested and focused, especially if the reader were a young person interested in learning how to play guitar. Participant 4 added that guitar playing is easier to learn through seeing and hearing than reading, but thought that the book might be useful for someone with previous experience with the subject, such as a music teacher preparing for a lesson. Participant 2 noted that since the flow of information is logical, it is easy to skip parts of the text if the reader is already somewhat familiar with the subject.

The layout was found to present some problems too. The participants found that the large amounts of text in somewhat small print made it difficult to use the book while holding a guitar. Participant 1 suggested that the problem could be resolved by using a music stand to raise the book to eye-level. Participants 2 and 3 found the layout to be heavy and cumbersome, while participants 1 and 4 thought that the chapter lengths and layout choices were appropriate and clear. The pictures were found to be good and helpful by participants 1, 2 and 3, while participant 4 was very critical of the pictures and thought they were unclear and poorly placed. Similarly, participants 1 and 2 thought that the pictures might initially be difficult to connect to the part of the text that they were referring to, but that all in all they were helpful. Participants 2, 3 and 4 commented that they were not familiar with some of the illustrations that were used alongside the pictures. They assumed these illustrations might be explained somewhere in a previous section of the text – however, they were not.

There were some overall improvements suggested by the participants. Three of them hoped for some audio–visual material to accompany the text. Participants 2 and 4 would have preferred tablature notation instead of the illustrations used in the text. Participant 1 suggested that the vibrato technique should be relocated to come before the part describing sliding, since the vibrato is referred to in that section. Participants 1 and 2 suggested that a larger font would be more useful when reading the book while holding a guitar. While most of the participants found the information flow to be logical, participant 4 thought that the information should be distributed differently: to concentrate on one skill level at a time, instead of extensive information on the single techniques. Most participants commented that a book would not be used as often as a form of learning material nowadays compared to online videos and tablatures.

A table detailing the most relevant findings for both tests is provided after the references in the Appendix.

6 DISCUSSION

The two methods, heuristic evaluation and usability testing, provided valuable information on the usability of the text. There were similarities between the results found by using the different methods, but some outcomes were also remarkably different. In addition, while the two tests might not be comparable as such, the results also portrayed a notable difference between what is valued in an evaluation done by language experts and an actual use situation with test users. Similarities found in the results of both tests included criticism on the pictures and layout, while the problems in the layout were emphasised more in the usability test. The book format was considered suitable for the purpose in both tests, but the test users seemed to emphasise the lack of audio–visual material more than the experts.

Perhaps the most notable difference was in how the translation's language quality was perceived differently between the tests. While the experts considered the language quality to be somewhat problematic in the heuristic evaluation, the usability tests' participants had no complaints: they found the language to be good and the text to be well written. This could imply that while the heuristic evaluation found the language to be at least somewhat problematic and deemed it would need revision, the actual use of the book might not seem to suffer because of the translation's language quality. Of course the different situations between the evaluation and the user test, as well as the background of the evaluators (language experts) and test users (non-experts) play a major role in the findings, but it is worth considering whether the language quality could in fact be seen as adequate for the book's actual users. This reflects a hot topic when discussing translation quality: the role of the reader versus that of the language expert. Some TS scholars such as House (1997:159) find the shift towards a more target-audience based approach as 'misguided'; language experts are preferred as those who determine the appropriate level of quality. However, as we have seen, actual test users can provide an interesting insight into the translation evaluation too, irrespective of whether or not their views correlate with the experts' notions of quality. Neither the heuristic evaluation nor the usability testing confirmed Valkonen's critique of 'impossible' language making the text difficult to understand. The evaluators commented on the language being 'awkward' and 'not at all terrific', but they agreed that at least the excerpt used in the evaluation manages to present its information well.

Further differences between the two tests include how the test users were less familiar with the symbols used alongside the pictures. This can be seen especially in the participants' wishes to include tablature notation. One problem found in the usability testing that did not appear in the heuristic evaluation is the amount of information, which was found to be too heavy to keep the reader interested. Indeed, while the experts found the text to be mainly suitable for beginners, the test users deemed it might not be interesting for a young person learning guitar, instead suggesting the book would be suitable as material for someone with previous knowledge on the subject – such as

participant 4's suggestion of a music teacher preparing a lesson. Some of the comments and suggested improvements made by the test users contradicted each other. It is indeed somewhat common in usability testing that different users make conflicting suggestions regarding improvements (Nielsen 1994:191).

To further evaluate the usability of the text, we can take Byrne's definition regarding the usability of a text: "When applied to texts usability measures the extent to which readers can read a text, understand its content and perform whatever task is required by the text quickly and accurately and the extent to which they find the experience difficult or easy" (Byrne 2012:201). If we break down the definition into its components, we can compare the results of the tests accordingly. The first part measures the extent to which the user can read the text. There were no major problems found in reading the text, but some issues were found in the layout. Similarly, the heuristic evaluation suggested that the language might cause some problems with reading, but the usability testing did not confirm this. Some test users and evaluators had concerns about the pictures, which can also hinder reading. The second part of the definition measures understanding of the content. Again, the results of the tests suggest that the text is understandable and the language does not interfere with understanding. However, the pictures could prove to be difficult to understand for some users.

Byrne's definition continues with being able to perform tasks required by the text quickly and accurately. As seen in the usability testing, all users were able to complete the task. The time taken to complete the test and be satisfied with the results varied from 10 to 25 minutes between users. The only problem in performing a technique accurately related to the horizontal vibrato, which most users were more familiar with performing vertically. It is noteworthy that the test users were all familiar with the techniques to some extent and as participant 4 put it, most players spend years trying to perfect these techniques, which are described here in only four pages. Regardless, the overall performance of the users could be described as quick and accurate. The final part of the definition refers to the extent to which the users find the experience difficult or easy. I would wish to broaden the definition to address user satisfaction here too, since it is mentioned in many definitions of usability, including Nielsen (1993:26), ISO 9241-11 (1998) and Suojanen et al. (2015:13). While the test users did not experience the use of the text as problematic themselves, most were concerned that the book might not be satisfying learning material for younger, inexperienced players. As an alternative, audio–visual learning material would have been preferred. Both the heuristic evaluation and usability testing would suggest that the book format is somewhat outdated as learning material for guitar playing. The heuristic evaluation suggested that the 'awkward' or somewhat difficult language makes the reading experience less satisfying, but this was not confirmed by the usability testing. However, the heavy layout and unclear use of graphics were found to affect the experience.

It could be argued, in fact, that most of the problems found in the usability testing were to do with user satisfaction, while the problems found in the heuristic evaluation were mainly concerned with language. This would correlate with Valkonen's criticism of the book's translation quality. However, the usability problems found in the tests were overall not as devastating as the criticism might suggest. The heuristic evaluation would suggest

that most of the problems were related to the two first heuristics, 'accessibility' and 'matching the real world', while problems concerning the other heuristics were less frequent. Problems related to compliance with target culture standards could be seen in the usability testing, where users found some illustrations to be unfamiliar and would have preferred that audio–visual material and tablature notation would accompany the text. The heuristics 'accuracy' and 'purposeful and ergonomic' were realised well, with only minor complaints from the evaluators. The usability standards of the 'information design' heuristic were deemed to be met well and 'user support' was sufficient, though the evaluators mainly found the text to be suitable for beginners. Interestingly, while the experts seemed to share the opinion that the text would suit less experienced users better, the test users found it might be off-putting for younger players with less experience and would be better suited for readers already familiar with the subject. It is interesting to note these instances where the test users occasionally took the role of an expert evaluator when discussing the material, and how they considered other potential users might feel about using the book.

It should be stressed, however, that most of the usability problems found in the tests were not necessarily connected to the translation itself. Some of the problems found in the expert evaluation could have been addressed through revising the translation, but there does not seem to be much the translators could have done in terms of layout and the pictures, since they apparently remained identical in both the original and the translation. Some of the problems, such as preferring audio–visual material to accompany the text, could be related to technological developments and changes in the self-learning methods of guitar players. Arguably, the expert evaluators could be seen to be closer to the book's original target audience age-wise, compared to the test users, who were in their 20s, considering that the book was first published in 1982.

As mentioned in section 3.2, the UCT model's application of usability methods iteratively within a translation process did not apply here, since the methods were used only to evaluate a finished translation. Further studies on the subject should take this point into consideration and look into which parts of a translation process could the methods be applied to most beneficially, as opposed to this type of end-of-the-line assessment.

7 CONCLUSION

Heuristic evaluation and usability testing demonstrated the potential for evaluating translations from a usability perspective. However, more work is certainly needed to ensure that usability methods are better suited to translation evaluation and revision. A usability-based approach could prove to be beneficial for translation processes and provide more specific information for improving the evaluated product. Here the methods were used to evaluate a finished product, not as a part of iterative product development. Indeed, usability methods would seem to be more beneficial when used to improve an unfinished product instead of finding problems only after they could be addressed. The heuristics used here would also benefit from refining. For instance, they did not address user satisfaction as such, which would be a central point of usability as seen in the various definitions presented above. There is also notable overlapping between the heuristics: For

example, under which heuristic should uncommon sentence structures and grammatical problems be reported? Similarly, the heuristics could be modified to better address such factors as satisfaction and efficiency. Refining the heuristics would also require field testing them in actual translation projects.

Looking at the heuristic evaluation and usability testing, the results of the two tests are naturally not comparable as such, since they are two different tests, using different methods for the same text. However, when looking at the process from a usability research perspective, they can be seen as complementing each other. It is indeed advisable that the more resource and cost-effective heuristic evaluation should be conducted earlier on in the product development process, and the more costly and time-consuming usability testing should be conducted later on in the process – preferably after the problems found in the heuristic evaluation have been addressed. It is also interesting to see how the tests produced different results where the language of the text was concerned. This might suggest that while the expert evaluators seem to pay more attention to the linguistic features, the actual users might not even see problems in the quality of the language, considering that they found it to be good and well written. This is worth considering when focusing on translation revision and evaluation, where most of the focus can often be on the linguistic features, thus overlooking other important issues that might be present in the text.

When looking at an application for using both heuristic evaluation and usability testing as a means of addressing usability problems of a translated text, it should be noted that expert evaluators and actual users might have different views on the evaluated product's usability. Thus the question of using the same methods to compare differences between the expert evaluators' and actual users' views would not be advisable, nor has it been the point of this study. However, the question of how different actors view the concept of quality is visibly present – the actual users might be happy with texts that are factually accurate but not linguistically sophisticated. Studying translation usability further could extract interesting insights into how to procure users with fit-for-purpose translations most beneficially. Further application of usability and UCT in Translation Studies and translator training, as well as in cooperation with members of the translation industry, could provide valuable information on how a usability-based approach might benefit the field of translation as a whole.

REFERENCES

Byrne, Jody. 2006. *Technical translation: Usability strategies for translating technical documentation*. Dordrecht: Springer.

Byrne, Jody. 2012. *Scientific and Technical Translation Explained*. Manchester: St. Jerome.

Denyer, Ralph. 1982. *The Guitar Handbook*. London: Dorling Kindersley.

Denyer, Ralph. 1982. *Suuri kitarakirja*. Translated by Saastamoinen, Ilpo, Jyrki Manninen, Juha Nuutinen, and Tapio Peltonen. Helsinki: WSOY.

Doherty, Stephen, & Sharon O'Brien. 2014. "Assessing the usability of raw machine translated output: A user-centered study using eye tracking." *International Journal of Human-Computer Interaction* 30 (1): 40-51. doi:10.1080/10447318.2013.802199

Gouadec, Daniel. 2007. *Translation as a Profession*. Amsterdam: John Benjamins.

House, Juliane. 1997. *Translation quality assessment: A model revisited*. Tübingen: Narr.

ISO 1998. *ISO 9241: Ergonomic requirements for office work with visual display terminals. Part 11: Guidance on usability.* International Organization for Standardation.

Lång, Juha. 2013. "Suomalaisten av-alan toimijoiden tekstityskonventioiden vertailua." *MikaEL: Kääntämisen ja tulkkauksen tutkimuksen symposiumin verkkojulkaisu vol. 7*. https://sktl-fi.directo.fi/@Bin/319343/Lang_MikaEL2013.pdf. Visited 2 February 2015.

Nielsen, Jacob. 1993. *Usability engineering*. San Diego, California: Morgan Kaufmann.

Nielsen, Jakob. 2012. "Usability 101: Introduction to Usability." http://www.nngroup.com/articles/usability-101-introduction-to-usability/ Visited 20 February 2016.

Purho, Vesa. 2000. "Heuristic inspection for documentation - 10 recommended documentation heuristics." *STC Usability SIG Newsletter* 6 (4), 24 September 2013.

Pym, Anthony. 2011. "Website localization." In *The Oxford Handbook of Translation Studies* K. Malmkjær & K. Windle (eds), 410-424. Oxford: Oxford University Press.

Salmi, Leena. 2003. *Documents multilingues pour logiciels et utilisabilité*. Turku: University of Turku.

Suojanen, Tytti, Kaisa Koskinen, and Tiina Tuominen. 2012. *Käyttäjäkeskeinen kääntäminen*. Tampere: University of Tampere.

Suojanen, Tytti, Kaisa Koskinen, and Tiina Tuominen. 2015. *User-centered translation*. London: Routledge.

Valkonen, Tero. 2000. "Kustantaja, kirjassani on virhe!" *Helsingin Sanomat*, weekly supplement *Nyt* 44/2000, 61.

APPENDIX

List of most relevant results

	Heuristic evaluation	**Usability testing**
Language of the translation	Overall language is not very good Finnish. Source language interference, some non-idiomatic expressions and unnecessarily noun-heavy expressions. Awkward features in word choices and syntax, but understandable language. Tone is 'motivating and inspiring, despite its awkwardness.'	Well-written, good Finnish language. Tone of the text is helpful and motivating. Educational style of the text is well-suited for the purpose.
Terminology	Most of the terms are clear and familiar to users having at least some level of musical experience. Some translated terms unfamiliar or amusing. Including *nauhaisku* (hammer on) and *nauhanyhtäisy* (pull off). Translation mainly consistent, except for terms used for bending: *venyttäminen* and *taivuttaminen*.	Overall familiar terminology, especially the English terms. *Nauhaisku* and *nauhanyhtäisy* unfamiliar and somewhat amusing translations for 'hammer on' and 'pull off' for most participants.
Technical and/or semantic errors	Factually accurate. Problems with grammar, such as missing or misplaced commas and uncommon sentence structures. Some typos.	Problems performing horizontal vibrato as described in the text. Horizontal vibrato described as the more common way of performing the technique, but one participant had never performed it that way before.
Focus of information	Information mostly effectively focused and relevant. Some unnecessary information, which could be more distracting than helpful.	Too much content to keep the reader interested. Especially younger readers might be put off by the amount of text.
Appropriate media	Book is a good interface. A–V material could be used to complement the book.	Book less typical learning material for test users compared to online tablatures and videos. Text could be complemented with A–V material. Tablature notation would have been preferred instead of the unfamiliar illustrations in the text.

Layout	Layout problematic. Fragmented impression. Pictures and their references unclear.	Layout heavy and cumbersome. Font should be larger. The book is difficult to read while holding a guitar. Too much text packed into a small space.