Juho Antti Suokas

NOTHING FREEZE-DRIED

TESTING UsABILITY Evaluation METHODS
WITH THE FinNISh TRANSLATION OF The GUITAR HANDBOOK

MA Thesis
2014
This study uses usability evaluation methods to analyse an excerpt of the Finnish translation of Ralph Denyer’s *The Guitar Handbook*. The book’s Finnish translation, *Suuri kitarakirja*, has been criticised for its language, which, according to Tero Valkonen (HS, *Nyt* 44/2000), is “impossible”. Inspired by Valkonen’s criticism, this aims to examine if the book is actually difficult to use and, as a secondary aim, to test usability evaluation methods in practice.

Usability is a relatively new concept in Translation Studies, although it has been a subject of Human-Computer Interaction studies from the 1980s. Usability focuses on the user of a product. The product should fit its purpose so that its use is effectively, efficiently and satisfyingly by specified users in a specified context. There is also a correlation between usability evaluation in Translation Studies and traditional translation quality assessment. In this study the usability of *Suuri kitarakirja* is evaluated by using heuristic expert evaluation and usability testing.

The expert evaluators are staff members of the School of Humanities of the University of Eastern Finland. They have expertise in language and experience with guitar playing. They were sent a questionnaire based on modified usability heuristics and the chosen excerpt of *Suuri kitarakirja*. The answers to the questionnaire are used as the basis for the heuristic evaluation. The participants in the usability testing are students of the University of Eastern Finland who play guitar but are not students or experts of language. The usability testing consisted of the participant user practising playing techniques from the chosen excerpt of the book, followed by an interview.

The results suggest that the language of the translation does not interfere much with its usability. The results of the expert evaluation suggest that the language is not very good Finnish, but it does not affect understanding. Similarly, the usability testing does not find many problems relating to the language. While most of the problems found in the expert evaluation are concentrated on the language of the translation, the problems found in the usability testing are mainly concentrated on the book’s layout and information presentation.

Interestingly, applying usability evaluation methods points out problems in the translation that are not necessarily addressed in traditional translation quality assessment. Usability evaluation would seem to present a new and interesting angle for evaluating translations and developing translation quality assessment models.

**Avainsanat – Keywords**

# Table of Contents

**ABSTRACT**  

**TABLE OF CONTENTS** ........................................................................................................ 1  

1. **INTRODUCTION** ........................................................................................................... 2  

2. **THEORETICAL BACKGROUND** .................................................................................... 5  

   2.1. **FUNCTIONALISM** .................................................................................................. 5  
   2.2 **USABILITY** .............................................................................................................. 8  
   2.3 **USABILITY EVALUATION** ..................................................................................... 11  
      2.3.1 Observation methods ............................................................................................. 14  
      2.3.2 Survey methods ..................................................................................................... 15  
      2.3.3 Heuristics ............................................................................................................. 18  
   2.4 **TRANSLATION QUALITY** ..................................................................................... 24  
      2.4.1 Defining quality ................................................................................................... 25  
      2.4.2 Quality Assessment ............................................................................................... 27  
      2.4.3 Quality Assessment Models .................................................................................. 29  

3. **MATERIAL & METHODS** ............................................................................................... 35  

   3.1 **THE GUITAR HANDBOOK & SUURI KITARAKIRJA** ............................................. 35  
   3.2 **METHODS** ............................................................................................................ 38  
      3.2.1 Expert evaluation .................................................................................................. 38  
      3.2.2 Heuristics ............................................................................................................. 39  
      3.2.3 Usability testing .................................................................................................... 45  

4. **RESULTS AND DISCUSSION** ..................................................................................... 48  

   4.1 **RESULTS OF EXPERT EVALUATION** ................................................................ 48  
      4.1.1 Matching real world & Accessibility ..................................................................... 48  
      4.1.2 Accuracy ............................................................................................................... 50  
      4.1.3 Purposeful and ergonomic .................................................................................... 51  
      4.1.4 User support ......................................................................................................... 52  
      4.1.5 Information design ................................................................................................ 53  
      4.1.6 Summary of expert evaluation .............................................................................. 54  
   4.2 **RESULTS OF USABILITY TESTING** ...................................................................... 55  
      4.2.1 Results of direct observation ................................................................................ 55  
      4.2.2 Results of interviews ............................................................................................. 57  
      4.2.3 Summary of usability testing ............................................................................... 60  
   4.3 **DISCUSSION** .......................................................................................................... 61  
      4.3.1 Discussion of the primary purpose ....................................................................... 61  
      4.3.2 Discussion of secondary purpose, methods and success of the tests ................... 65  

5. **CONCLUSION** .............................................................................................................. 72  

**REFERENCES** ................................................................................................................ 74  

   MATERIAL .......................................................................................................................... 74  
   LITERATURE ........................................................................................................................ 74  
   INTERNET REFERENCES .................................................................................................... 79  

**APPENDIX**  

   Text excerpt  

Finnish Summary
1. Introduction

“The Finnish reader must fight their way through impossible language to get to the point.”

This beginning sentence is translated from the article “Kustantaja, kirjassani on virhe!” by Tero Valkonen, published in Nyt (44/2000), the weekly addition to Helsingin Sanomat. Valkonen writes about the mistakes and poor quality of language in translated books and presents the Finnish translation of Ralph Denyer’s The Guitar Handbook (1982) as a case in point. The Finnish version is titled Suuri kitarakirja (1982), translated by Ilpo Saastamoinen, Juha Nuutinen, Tapio Peltonen and Jyrki Manninen. While Valkonen gives praise to the original work, he claims that the translated version has been “severely damaged” by the translators and made “all but unreadable” because it contains “every possible translation mistake there is”, and that the book could be used as educational material of how not to do translations. Language professionals – especially translators (including Valkonen himself) – can, indeed, become extremely critical of the language they read or write. It is their profession, after all. However, using Suuri kitarakirja as an example of bad translation would be regrettable from the point of view of its author and translators, because the book aims to be educational material for various guitar-related topics instead. Yet what about the readers’ point of view? Is the language of the translation troublesome for someone using the book to practise guitar playing? Perhaps most of the intended audience would not be as troubled by the language as Valkonen.

The key point to consider here is the sentence in the beginning of this chapter. If the language makes the book problematic to use, this could be seen as a usability problem in the translation. Usability is a fairly new concept in Translation Studies (later TS) and, inspired by Valkonen’s

---

1 All comments by Valkonen in this study are translated by JS.
criticism, this study attempts to apply usability strategies in order to have a model on which to base the evaluation of *Suuri kitarakirja*. Usability can be seen to correspond in part to the hot topics of translation quality (TQ) and translation quality assessment (TQA). Since TQA can be a notoriously difficult subject to tackle, this study attempts to approach the subject from a usability perspective. Quality will be discussed and we shall compare some usability evaluation methods with TQA principles, but quality itself is not in the focus of this study.

The primary purpose of this study is to analyse an excerpt of *Suuri kitarakirja* with usability methods. The secondary purpose is to test the application of usability evaluation methods and comparing them with theoretical TQA models. There are two main stages to this study: The first stage is to examine the concepts and theories behind usability and translation quality assessment and to find a suitable model for their application. The second stage is to apply these methods in a case study of the Finnish translation of *The Guitar Handbook*. This study uses expert evaluation and usability testing as the chosen methods of assessment. The evaluation is performed summatively (see Chapter 2.3) and it focuses on usability and adequacy of the evaluated product instead of textual equivalence of the source text and target text.

The title for this study draws from a sentence from *Suuri kitarakirja*, presented by Valkonen as a prime example of the Finnish translation. On page 28 of *The Guitar Handbook* (1982, same pages in the original and translation), Frank Zappa scorns the playing style of Elvis Presley’s session guitarists, such as Scotty Moore and James Burton. Instead, Zappa suggests examples of better players, for instance Johnny Watson or Guitar Slim. Here Zappa compares the two playing styles by stating, "[t]hat's a guitar solo, nothing freeze-dried." Valkonen ironically states that here the translator (Manninen in this case) has really “bent over backwards” by translating Zappa’s comment as “[n]e ovat kitarasooloja eivätkä mitään pystyynjäätyneitä kuivuuksia.”
This thesis is structured into five chapters. Chapter 1 is this introduction. Chapter 2 discusses the theoretical background, including functionalist translation theory, usability, usability evaluation and translation quality. Chapter 3 introduces the material used for the testing as well as the chosen methods. In Chapter 4, the results of the expert evaluation and usability testing are presented and discussed. The methods, their application and the success of the tests are also discussed in Chapter 4. Chapter 5 is the conclusion.
2. Theoretical background

In this study, we shall examine how usability research methods can be applied to examine translations and translation quality. In this section we shall focus on usability and look at translation quality, but we begin by examining functional translation theory, which provides the general framework for the methods used in this study.

2.1. Functionalism

The term ‘functionalism’ in Translation Studies refers to theoretical approaches in which the most important assessment criterion for any translation is the function or purpose of the target text. In comparison with non-functionalistic approaches, these do not focus extensively on the linguistic ‘equivalence’ of the source text (ST) and target text (TT), but instead place emphasis on the translator and the users of the translation (Schäffner, 1998; Hönig, 1998).

Functionalist approaches are seen to have developed from Hans Vermeer's skopos theory of translation. The theory focuses on the skopos – the purpose – of translations instead of concentrating on their linguistic features. Translating is here seen as a sociocultural human action and thus the translation should address the needs of its recipients. The theory was developed in Germany in the late 1970s, distancing itself from previous translation theories that were often focused on literary translation, while the skopos theory addressed the translation of non-literary translations and their cultural contexts. Vermeer's idea was that translating is human action, which in turn is determined by the purpose of the action. Thus, the translation (action) is a function of its purpose (skopos). In practice, this would suggest that the translation's requirements are largely defined by the initiator, or client, as well as the
constraints of the TT reader’s situational and cultural background. This was a major departure from equivalence-based translation theories, where the translation is defined by factors such as the source text’s linguistic functions or effects on the reader (Vermeer, 1996; Schäffner, 1998: 235–238).

For the purpose of this study, the notion of equivalence will not be used as a basis for translation quality – although it has often been used as such. Even the term ‘equivalence’ itself seems to be a controversial one, with multiple definitions from different theorists. Dorothy Kenny (1998: 77–78) claims that some theorists use equivalence as the key component of defining translation while others might reject it completely. In addition, Kenny points out that most definitions of equivalence are actually circular: “equivalence is supposed to define translation, and translation, in turn, defines equivalence” (1998: 77). However, I must point out that Suojanen et al. (2012: 43–44) see Eugene Nida’s concept of dynamic equivalence to correlate with the usability aspects we will be examining later. Nida was a linguist and translation theorist who focused on Bible translation and created the concepts of dynamic and formal equivalence. While formal equivalence emphasises the form and contents of the message, dynamic equivalence focuses on translation as dynamic communication that is bound to cultural and social contexts. When considering the active role of the recipient, or the reader of a translation, dynamic equivalence focuses on conveying a similar effect on the reader instead of merely translating the words. Thus, the dynamic effects between the TT and its reader should be similar to the ST and its reader (Suojanen et al. 2012: 43–49).

In functionalist translation theories, adequacy is seen to be the important factor on the basis of which to assess translations (see e.g. Vehmas-Lehto, 1989: 16–17). Adequacy, for the purpose of this study, is defined along the lines of how Reiss & Vermeer (1986), Vehmas-Lehto (1989)
and Nord (1997) have presented it: adequacy is seen as a quality of the product, which serves the purpose of the desired communication act.

Within the framework of Skopostheorie, ‘adequacy’ refers to the qualities of a target text with regard to the translation brief: the translation should be ‘adequate to’ the requirements of the brief.

(Nord 1997: 35)

It should be noted that the term ‘adequacy’ is not without problems in Translation Studies. Adequacy can be found in Gideon Toury’s ‘Descriptive Translation Studies’, or DTS, where the term is used in quite a different setting, describing the norms of adequacy vs. acceptability. Toury’s norms can be seen as a regular set of patterns and strategies used in the decision making process of translating – either prior to or during the actual translation process. In DTS adequacy is seen as an initial norm, which is a choice of adhering to the norms of the source text language and culture, while acceptability is in turn seen as adherence to the norms of the target language and culture (Baker, 1998: 163–165).

Functionalism has not been without its criticism and problems. Hans G. Hönig (1998: 14) points out when regarding functionalist translation theory that “functionalism begs the question of supposed reader's response.” However, according to Colina (2009: 238), “reader-response testing is time-consuming and difficult to apply to actual translations.”

Of course, functionalism cannot be seen as an all-inclusive translation theory. It should be seen as one approach among others. However, functionalism in my opinion could be seen as complying with the evolution of the field and the world surrounding it. Indeed, Translation Studies should not be seen as an academic discipline entirely separate from the actual practice, since translation is mainly a practice-oriented field. To give a skopos-inspired example, it
would not serve an underpaid translator who is being paid by the piece to spend countless hours of work on minor adjustments on such matters as the quality of language.

2.2 Usability

Usability, as seen here, has its roots in Human-Computer Interaction (HCI) studies (Suojanen et al., 2012: 15). Jacob Nielsen – a well known HCI usability expert – defines usability as “a quality attribute that assesses how easy user interfaces are to use” and adds that “[t]he word ‘usability’ also refers to methods for improving ease-of-use during the design process” (Nielsen, 2012). According to Nielsen usability is thus not restricted to the assessment of certain qualities, but it includes the aspect of improving these qualities as well. In addition, usability is defined as an ISO standard (ISO 9241-11) as “the 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.” Nielsen also mentions “utility”, which describes whether the product “provides the features you need” and concludes that a product’s “usefulness” is a combination of usability and utility (Nielsen, 2012).

We can see that the incorporation of usability into a translation process could benefit the overall ‘quality’ of the final translation (see e.g. Byrne, 2006: ix). In addition, evaluating usability can be seen to correspond with evaluating quality – at least certain aspects of quality. It must also be taken into consideration that while usability testing is commonly present in the production process, in this study it is used as a means of evaluating an already translated product.
Nielsen presents five quality components that define usability:

- learnability (ease of use when the product is first encountered),
- efficiency (the users’ performance speed with the product),
- memorability (ease of use when returning to the product after a period of time),
- errors (the number, severity and ease of recovery from errors users make with the product) and
- satisfaction (how pleasant is the product for the user).

(Nielsen, 2012.)

Nielsen’s focus is mainly on internet and intranet user interface designs. He exemplifies usability as an essential part of web-page design by suggesting that if a page is not easy to use and its information is not easily accessible, visitors will leave the page in favour of a better designed one (ibid.). However, usability is not restricted to merely Human-Computer Interactions. It has been applied to products and services, including texts and translations, as can be seen in works by such authors as Byrne (2008, 2012), Suojanen, Koskinen & Tuominen (2012) and Purho (2000).

Usability has become an increasingly popular subject in Translation Studies. The term ‘user’ has not been commonly used in Translation Studies, but it can be found in the works of such authors as Hönig (1998), Colina (2008, 2009) and Pym (2010). More recently the terms ‘user’ and ‘usability’ have been integral in the works of Jody Byrne (2006, 2012). Byrne’s focus is on technical translation and how usability strategies can be used to improve their quality. To give an example, Byrne defines the 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.

Byrne (2012: 201)

Byrne has distilled his definition from various sources, such as the ISO 9241-11 standard – which covers HCI ergonomics – and writings by authors such as Dumas & Redish (1999, in Byrne, 2006: 97–98). I wish to point out three aspects of Byrne’s definition of usability: 1) the focus is on the readers/users of the text, 2) the readers are using the text to perform a task, and 3) the experience is defined by the users themselves.

Byrne's focus is on technical translation, but usability should be seen widely applicable to other forms of translation as well. Käyttäjäkeskeinen kääntäminen (2012) by Suojanen, Koskinen & Tuominen is, according to the authors, taking off where Byrne has finished, examining usability in translation on a larger scale. While usability is seen to benefit mostly instructive texts (Byrne, 2006: 255; Suojanen et al. 2012: 32–33), the authors broaden the scope to cover other types of translations as well. The authors have a functionalist viewpoint, emphasising that translation is instrumental, it is always needed to perform a purpose (Suojanen et al. 2012: 12). In addition to Byrne's aforementioned definition of usability, Suojanen et al. see usability as user and context specific, emphasising both social aspects – such as accessibility and social acceptability – and user experience aspects – such as personal intuition and affective factors (2012: 15–20). Suojanen et al. offer what they call User-centered translation (UCT), a model or toolkit which incorporates users and usability methods into the translation process.
2.3 Usability evaluation

The evaluation of usability can be done either formatively or summatively (Byrne, 2006: 177–178). Formative evaluation takes place during the design and development of a product. An example of this would include a translation project that employs UCT methods in the translation process to improve the usability (and quality) of the final product. In contrast, a summative evaluation takes place after the product is finished. This study is an example of a summative evaluation, which evaluates the usability of a finished product.

There are many different methods of evaluating and testing usability. Noticeably, with regard to the usability of texts, especially translations, empirical methods are preferred (see e.g. Byrne, 2006: 179–181; Suojanen et al. 2012: 69–73). Byrne divides empirical usability evaluation into two categories: methods which include users and methods which do not (2006: 180). He suggests that those evaluation methods which involve actual users produce more relevant information. Accordingly, Nielsen (1997) also presents usability testing with users as the most basic and useful method of studying usability. These user-based methods include various different testing possibilities, including methods already in use in Translation Studies, for instance eye-tracking, thinking aloud and the use of interviews and questionnaires. In addition, we shall examine heuristic evaluation, which does not necessarily involve actual users, but in which the evaluators are considered experts.²

Before conducting a usability test, careful planning is required. Rubin & Chisnell (2008: 67) point out the following parts, which are most commonly included in all user-based usability test plans:

² I use the expression ‘not necessarily’ here, since in some cases these experts could be seen as a part of the target user group too.
Purpose, goals and objectives of the test
Research questions
Participant characteristics
Method (test design)
Task list
Test environment, equipment and logistics
Test moderator role
Data to be collected and evaluation measures
Report contents and presentation

Rubin & Chisnell (2008: 67)

As can be seen from the list, the first step is to justify the usability testing, to decide whether it fits the purpose or not. The second part, research questions, is according to Rubin & Chisnell (2008: 69) the most important one, since this dictates the rest of the testing by defining the questions the test wishes to answer. Rubin & Chisnell maintain that this is equally important in experimental, less structured tests, since the test conductors need to be aware of what they wish to learn from the test (ibid.).

The third part, participant characteristics, defines the test group. The test group should reflect actual users of the product being tested, which would require knowledge of the product's users or specific user profiling to be able to select suitable test participants (Byrne, 2006: 194–195). The number of participants is important, since too few participants do not produce sufficiently accurate results. Rubin & Chisnell (2008: 72) suggest using 10–12 participants per condition when conducting a formal usability test. However, less formal usability testing can be conducted using 4–5 participants to represent the intended audience, since such a group can find out around 80 per cent of the test product's usability problems (Rubin & Chisnell, 2008: 72; Suojanen et al. 2012: 71; Nielsen, 2000). In addition, Nielsen (ibid.) suggests using no more than five users and instead conducting as many tests as possible. However, Rubin & Chisnell point out that the remaining 20 per cent of usability problems might be important for the
product. A larger group is suggested especially if the moderator of the test does not have much experience; this gives more possibilities to practice moderating skills and reduces the risk of missing important problems (2008: 72–73).

The fourth part, the method of the test describes how the test will progress, what to expect from the moment when the participants arrive to when they leave. The task list describes what will happen during the test, it should include tasks that correspond with the actual use of the product/text being tested. Suojanen et al. (2012: 71) point out that the language used to give these tasks should be unambiguous, direct and natural, and it should not manipulate the user towards certain outcomes. In the light of previous research, Byrne (2006: 202) also suggests that the material used for the testing should be edited for “typographical errors, style inconsistencies, grammatical or punctuational errors.” Rubin & Chisnell (2008: 80) recommend defining beforehand what counts as a successful completion of a task, since there might be opposing views on this matter. The test environment should resemble or simulate an actual use environment for the product and user, the equipment described in this context includes only those used by the test group, not those used by the moderators. The seventh item on the list, the role of the moderator, is also important to define beforehand, since the moderators are the only ones who should interfere with the progress of the test situations (Rubin & Chisnell, 2008: 87–88). The data being collected should be based on the second part of planning, the research questions, and it is also dictated by the equipment being used to gather data. This can include measured variables such as error rates, tracked eye movements and time taken to complete tasks, but also immeasurable factors such as data gathered by

---

3 The test conductors can also be separated as facilitators and moderators. In such a division, the facilitator is seen as someone who controls the progress of the test, while the moderators are present conducting the test, but do not interact with the users. This division can be useful if a test has many conductors with different roles (observers, interviewers etc.). However, in order to keep terminology less complex, I shall keep to the term ‘moderator’ here. (For more terminology, see e.g. http://www.usabilitybok.org/glossary.)
questionnaires or interviews. The final part, report contents and presentation, includes a summary of the test report and how the results will be communicated further on. (Rubin & Chisnell, 2008: 67–91; Suojanen et al. 2012: 69–72.)

Now we shall take a look at some methods of gathering data in user-based testing, presented by Byrne (2006) and Suojanen et al. (2012).

### 2.3.1 Observation methods

Byrne suggest user observation as one of the best ways of gathering data. The observation can be carried out in a specifically created setting (laboratory) or in the users’ natural environment (field study), either directly or indirectly. Direct observation requires the users to preform the task while being watched by one or more observers, who gather data from the test. This method is useful, since it is informal and immediate in its nature. However, the presence of one or more observers might affect the users’ performance and the data gathered relies on the observer’s attention (Byrne, 2006: 181–182).

With indirect observation there is no observer present while the users preform the task, but their actions are recorded. Recording methods can include video cameras, software logging or eye-tracking. Video recordings can be used in place of direct observation, for the presence of an observer is not affecting the situation, and multiple cameras can be used to record different events. The benefit of video recording as opposed to direct observation is also the possibility to review the recorded material and return to specific occurrences, which might have been missed before. Software logging records computer interactions, which are commonly either time-stamped keypresseses (which keys are pressed and for how long) or interaction logging, which records the complete interaction during the test. (Byrne, 2006: 182–184.) Eye-tracking is
carried out with specific equipment that records the user’s eye-movements. This can be used to gather information on what the user has been observing and focusing on during the test (Suojanen et al. 2012: 75–76). It has been used in Translation Studies in many audio-visual reception studies, such as Lång et al. (2013).

Qualitative data can also be gathered by having the users vocalise their thoughts and ideas while performing the task – this is known as thinking aloud. These instances are usually recorded and transcribed, the transcriptions are known as think-aloud protocols, or TAPs in short. Thinking aloud has been borrowed into Translation Studies from cognitive psychology and it has been applied especially in translation process research (Jääskeläinen, 2010: 371–372). Thinking aloud can provide a wealth of useful information of the user interaction and cognitive processes involved, it is also seen as cost-efficient and relatively simple to use (Byrne, 2006: 185, Suojanen et al. 2012: 75). However, there are some drawbacks in using the method. To give some examples: vocalising one’s thoughts can take up much cognitive resources and affect the process, only conscious thoughts can be verbalised – leaving out automated processes and subconscious thought – and the theoretical basis for the method has been questioned (Byrne, 2006: 185; Jääskeläinen, 1998: 266–267; Suojanen et al., 2012: 73–75). Byrne also suggests that the extra cognitive effort required during the task can hinder users' performances and make TAPs a less accurate means of gathering usability information (2006: 201–202).

2.3.2 Survey methods

Another way of gathering data are survey methods, such as questionnaires and interviews. While observation methods are extremely useful when evaluating Nielsen's usability quality components such as learnability, efficiency and errors, survey methods address most of all Nielsen’s fifth component, satisfaction. Survey methods can provide qualitative information
especially on what the users want from the product, which would not present itself in observational usability testing (Nielsen, 1997). Byrne (2006: 187) suggests that objective information gathered by observational methods is not enough, for users’ subjective opinions are a very important part of usability and can point out “problems which may not have been anticipated by the designers or evaluators” (ibid.).

Interviews can be structured, flexible or semi-structured, according to what kind of information the interviewer wishes to achieve. Structured interviews include predetermined questions, which are asked in a fixed order. The benefit of structured interviews, as stated by Byrne, is being in control of the gathered data and the simplicity of its analysis. Flexible interviews, on the other hand, do not follow a strict pattern, but a list of topics the interviewer may or may not include in the discussions. The interviewer is free to follow up on interesting new topics, but the data gathered can prove to be more difficult to analyse than that gathered via a structured interview. Semi-structured interviews are a mix between the two aforementioned types, using a set of predetermined questions which the interviewer is free to use – or not to use – as they please. According to Byrne, the downside of the more flexible interview types is that they require an experienced interviewer, as well as the difficulty in analysing the less structured data (Byrne, 2006: 186–188).

Another interview-based survey method is using focus groups. Suojanen et al. (2012: 77, translation JS) define the focus group method as “a semi-structured group interview, administered by an interviewer or moderator.” The composition of the group and its context is important when using this method, for they have an effect on the data produced. As with other survey methods, it is suggested that focus groups are used in connection with other methods, such as user testing (Nielsen, 1997; Suojanen et al. 2012: 77–78).
Compared to interviews, questionnaires are easier to administer and analyse, but they require careful planning to produce proper results. They can be either self-administered or interviewer-administered. Self-administered questionnaires can reach large audiences for they do not require an interviewer, instead they are completed by the users themselves. Self-administered questionnaires require careful designing, for the user might misunderstand the questions or be misled by the wordings. In interviewer-administered questionnaires an interviewer asks the questions and gathers the data. The benefit of interviewer-administered questionnaires, as opposed to self-administered ones, is the higher response rate and possibility to control the process and clarify questions (Byrne, 2006: 188–190).

When using survey methods, the question types can be divided into three broad categories, presented by Byrne (2006: 189–190): factual, opinion and attitude questions. Factual questions represent actual facts about the users, such as which products they have experience with and how long they have been using said products. Opinion questions ask what the users feel; example questions could include whether the user prefers one product over another. Attitude questions aim to find out users’ attitudes towards the product being used. These can include question topics such as impressions of being efficient with the product, whether the user likes the product or how helpful and easy to learn the product seems. The questions can be presented as open or closed; open questions are answered in the user’s own words, while closed questions are answered by choosing from a set group of predetermined answers (ibid: 190).

It should be kept in mind, however, that usability tests are always created, artificial situations, and as such cannot be completely relied on to point out all usability problems (Suojanen et al., 2012: 72).
2.3.3 Heuristics

As an option to – or in addition to – testing with users, usability can be evaluated by heuristic evaluation, or expert evaluation. Nielsen (1995a) describes heuristics – as used in user interface design – as a method for testing usability, conducted by "a small set of evaluators [who] examine the interface and judge its compliance with recognized usability principles (the 'heuristics')." According to Nielsen, one single person is not enough for conducting a heuristic evaluation. The common recommendation is that heuristics, much like usability testing, should be done using 3–5 evaluators (Nielsen, 1995a; Byrne, 2006: 196; Suojanen et al. 2012: 101). These evaluators can be usability experts, novices or experts with knowledge on both usability and the evaluated product (Suojanen et al. 2012: 101). First the evaluators go through the product individually by using a list of recognised usability principles, or heuristics. They should not be allowed to communicate before the individual evaluations are finished (Nielsen, 1995a). Suojanen et al. (2012: 100) suggest the evaluators discuss their findings together after the individual evaluations and produce a report according to their findings; however Nielsen (1995a) proposes that the conductor of the evaluation can gather individual written reports from each evaluator or work as an observer, who monitors the evaluation situation and gathers data from the evaluators.
Nielsen presents a list of ten usability heuristics for use in interface design. These are as follows:

(1) Visibility of system status  
(2) Match between system and the real world  
(3) User control and freedom  
(4) Consistency and standards  
(5) Error prevention  
(6) Recognition rather than recall  
(7) Flexibility and efficiency of use  
(8) Aesthetic and minimalist design  
(9) Help users recognize, diagnose, and recover from errors  
(10) Help and documentation

(Nielsen, 1995b)

To clarify this list, I present Byrne's paraphrased version of these heuristics:

- Use simple and natural language.
- Say only what is necessary.
- Present the information in a logical way.
- Speak the users' language – use familiar words and concepts.
- Minimise the users' memory load.
- Be consistent.
- Provide feedback and tell users what is happening.
- Provide clearly marked exits to allow users escape from unintended or unwanted situations.
- Provide shortcuts for frequent actions and users.
- Provide clear, specific error messages.
- Where possible, prevent errors by limiting the number of available options or choices.
- Provide clear, complete help, instructions and documentation.

(Byrne, 2006: 162)

From Byrne's definitions, we can see how these heuristics could be beneficially applied to designing or analysing texts. Byrne (2006: 163) also elaborates how these heuristic principles can be worked into context-specific usability guidelines, such as: "Always phrase instructions consistently ... Avoid excessively long sentences ... Only use approved terminology ... Use the same formulations and constructions for sentences ... Avoid confusing verb tenses."
In addition, Nielsen presents a severity rating system for usability problems (1995c), which he suggests should be sent to the evaluators only after the initial heuristic evaluation. Using the rating system, the evaluators assess the usability problems on a scale of 0 to 4 accordingly:

- 0 = I don't agree that this is a usability problem at all
- 1 = Cosmetic problem only: need not be fixed unless extra time is available on project
- 2 = Minor usability problem: fixing this should be given low priority
- 3 = Major usability problem: important to fix, so should be given high priority
- 4 = Usability catastrophe: imperative to fix this before product can be released (Nielsen, 1995c.)

Nielsen believes that the ratings of a single evaluator are not reliable enough and suggests using the mean severity rating of at least three evaluators when applying these ratings (1995c).

Purho (2000) has also taken the idea of Nielsen's heuristics further and gathered a similar list for evaluating the usability of technical documentation. Akin to Nielsen's list, Purho's list consists of ten usability heuristics presented below, with explanations after statements when deemed necessary:

1. Match between documentation and the real world
   [The language is familiar to the user, documentation is logical.]
2. Match between documentation and the product
   [Same terminology used in product and documentation.]
3. Purposeful documentation
   [Clear intended use for each document and media fit for purpose.]
4. Support for different users
5. Effective information design
   [Information easy to find and understand. Purposeful graphics and use of language.]
6. Support for various methods for searching Information
   [Layout, index and form should support different users' information search methods.]
7. Task orientation
   [Documentation structured around independent user tasks.]
8. Troubleshooting
9. Consistency and standards
   [Consistent terminology and structure in each document. No unnecessary overlapping]
10. Help on using documentation
    (Purho, 2000; comments by JS)
Nielsen’s and Purho’s heuristics have been applied and tested in various Finnish pro gradu theses. Here I shall look at two of these by Reinikainen (2008) and Hämäläinen (2008).

Reinikainen (2008) applies Nielsen’s heuristic analysis and a focus group interview to evaluate the usability of Dungeons & Dragons 3.5 role-playing game. He relates the game’s rules to a computer user interface and his results show that role-playing games can indeed be analysed using Nielsen’s basic usability principles (Reinikainen, 2008: 76–77). Reinikainen suggests that the complexity of the game’s rules, as well as the foreign language (Finnish-speaking test group and rules in English), hinder usability – especially the immersion experience of the players is affected.

Hämäläinen’s (2008) study focuses on using Purho’s heuristics to evaluate the usability of the English documentation of Apple’s 5th generation iPod. In addition to evaluating the documentation, Hämäläinen also comments on the applicability of Purho’s heuristics as a method of evaluating usability. He proposes that the overall usability of the material is uneven (Hämäläinen, 2008: 82). In addition, Hämäläinen points out that Purho’s heuristics are well applicable to testing user documentation, but there is room for improvement. He suggests that the heuristics could be modified to account more for “predictability, memorability, error prevention, and user control and freedom” (2008: 83).

It should be taken into consideration that both Reinikainen and Hämäläinen did the heuristic evaluations themselves instead of using the recommended 3–5 expert evaluators. However, for the purposes of their studies – especially when considering Reinikainen’s use of a focus group and Hämäläinen’s examination of Purho’s heuristics – this can be seen as sufficient.
In Translation Studies, as a term ‘heuristics’ is not commonly used. However, Suojanen et al. compare heuristics with the quality assessment models that are present in most – if not all – translation projects (2012: 109). The authors suggest Gouadec's (2007) quality assessment principles as a list of noteworthy translation heuristics. Similarly, although Gouadec does not use the word ‘usability’, his definitions of a quality translation (2007: 6–8) are seen to correspond with Nielsen’s (1995a) usability factors. In these definitions Gouadec proposes that the final translated product must comply with "a) the client's aims and objectives" and/or "b) the user's needs or requirements" and at all times "c) the usage, standards and conventions applicable" (2007: 5). The definition of a quality translation according to Gouadec is as follows:

- **Accurate** – the content of the translation should be true to facts, ideally it should have no factual, technical or semantic errors (although this is rarely possible).

- **Meaningful** – the message, including concepts and connotations in the translation have to be meaningful in the target language and culture.

- **Accessible** – the message must be clearly understandable; the translation is adapted to fit the end-user; the translation must be readable, coherent, logical and well-written.

- **Effective and ergonomic** – the translation must effectively communicate its message and fulfil its function.

- **Compliant with any applicable constraint** – these constraints can be for instance legal, organizational, physical, functional or related to the target communities' linguistic and cultural standards and usages.
- **Compatible with the defence of the client's or work provider's interests** – the translator works for their client, the translation achieves its desired effects.

- **Economically viable** – efficient and cost effective.

  (Gouadec, 2007: 6–8)

It is noteworthy that, unlike Nielsen's list, Gouadec's definitions include the perspective of the client or work provider, which is an integral feature of translations. In addition, Gouadec's focus is not on the equivalence between the ST and the TT, but rather a quality translation is seen to fulfil its function for the user as well as its function in the target language. Again, there is a clear correlation to usability in these principles. For instance one can see much overlapping between Gouadec’s principles and Nielsen’s heuristics and usability quality components, especially in regard to efficiency, errors and consistency. Gouadec does, however, suggest that his definitions are not necessarily applicable to literary translations (2007:5). The applicability of usability methods to literary translation is also discussed by Suojanen, Koskinen & Tuominen (2012: 33–34).

Now that we have touched upon the issue of translation quality and quality assessment, we shall examine some of these aspects more closely, in order to see how they can be used in accordance with usability methods to evaluate translations.
2.4 Translation quality

As pointed out by Byrne (2006) and Suojanen et al. (2012), usability methods in translation have a clear connection to translation quality, which remains an undoubtedly hot topic in TS. Recent developments in the industry have raised questions of whether the state of translation quality is in decline (Vitikainen, 2013). Noticeably the increasing use of non-professional translation, such as online crowdsourcing (Susam-Saraeva & Peréz-González, 2012), and the current challenges that translators, especially those in the AV industry, are facing in Finland have been seen as a concern when regarding, not only the quality of translations, but the future status of and appreciation for the whole profession as well (the Finnish Association of Translators and Interpreters, 2012).

Quality is without doubt an important part of translation studies and translator education. For most translators quality is a matter of professional pride. This can be seen, for instance, in the use of pseudonyms when translators do not want their own name to be linked to their work. To give an example, an often used pseudonym among Finnish literary translators from the late 1940s up until 2000 was Lea Karvonen, which was used for instance when one was not happy with the quality of the translation or when working with less prestigious literary works (Kujamäki, 2007). As a more recent example, Finnish AV translators have been leaving their names completely out of some works to avoid being linked with poor quality subtitles (Vitikainen, 2013).

It must also be taken into consideration that the translation industry is not composed of only trained, professional translators (see e.g. Susam-Saraeva & Perez-González, 2012); thus we cannot see quality as merely a result of an acquired translator training. Quality has to be seen on a larger scale. In this section, translation quality (TQ) and translation quality assessment (TQA) will be analysed. Some methods of TQA are also explored.
2.4.1 Defining quality

What is quality? The question is not an easy one. Defining quality is difficult, for it is often an elusive, and multi-layered term which depends on context, in this case the quality of translations. Quality must be comprehended before it can be measured, as Abdallah (2007) points out. Thus we must examine what is meant by the term for the purpose of this study before we can proceed to assessing quality. Here quality will be examined and viewed in terms functionalism and usability.

The quality management systems standard ISO 9000:2005 defines ‘quality’ as “degree to which a set of inherent characteristics (3.5.1) fulfils requirements (3.1.2).” This quite simple explanation is as good a starting point as any. In the past, a good quality translation has often been seen as an “accurate, correct, precise, faithful, or true reproduction of the ST” (Schäffner, 1998: 1). However, there has been a shift towards seeing translation as text production, not reproduction. As Schäffner (ibid.) points out, the “basic tenet is that we do not translate words or grammatical structures, but texts as communicative occurrences.” This can be seen as a move towards a more functionalist approach in translation quality. This is remarkably present in the aforementioned quality principles, presented by Gouadec (2007: 6–8).

For this study, the main focus is on the quality and usability of a final product – in this case the translated Suuri kitarakirja –, but we shall also briefly examine other aspects of translation quality in order to achieve a broader sense of the term. For instance, Abdallah claims that the definition of ‘quality’ cannot be limited only to a high standard of language in the final translation. She views translation quality in the context of ‘Total Quality’, which involves three dimensions: “product quality, process quality and collective quality” (Abdallah, 2012: 5).
Product quality is the dimension of quality visible to the end user – the reader. Process quality is seen as how the translator works and with what equipment. Collective or social quality involves the questions of who works and under which circumstances. The concept of total quality can be seen to include an ethical perspective as well (Abdallah, 2007). Although the focus here is on the usability and quality of the final product, the process and social dimensions should be kept in mind too, especially in view of the professionalism or non-professionalism of the translators.

Another question to consider is who defines quality. Abdallah (2007) proposes that quality is not defined by language experts, instead it is based on the needs of large companies. This view on quality would include various extralingual aspects such as cost-efficiency, customer satisfaction and fast delivery times. The quality of translations can also be seen as a matter of reputation and corporate image – for instance, poorly translated commercial websites could seem unappealing to target audiences and cause damage to the company’s image. However, while many translation theories in the past have focused on defining quality along the lines of linguistic equivalence and adherence to cultural norms, translation quality should not, in my opinion, be seen as merely a feature assigned by language experts, but as complying with the needs of the client and the user, and fulfilling the skopos of the translation. As can be seen in Gouadec’s definition, a quality translation is one which takes into account the client’s interests and is efficient and cost effective (2007: 8).
2.4.2 Quality Assessment

Although we have chosen usability evaluation as a means of quality assessment for this study, in this section we shall take a look at the complex world of translation quality assessment and attempt to draw some parallels between usability evaluation and translation quality assessment.

Hönig (1998) points out his view of why translation quality assessment is necessary as follows:

> Users need it because they want to know whether they can trust the translators and rely on the quality of their products.

> Professional translators need it because there are so many amateur translators who work for very little money that professional translators will only be able to sell their products if there is some proof of the superior quality of their work.

> Translatological research needs it because if it does not want to become academic and marginal in the eyes of practising translators it must establish criteria for quality control and assessment.

> Trainee translators need it because otherwise they will not know how to systematically improve the quality of their work.

(Hönig, 1998: 15)

There are, however, various opposing views as to how translation quality should be assessed. The evaluation of usability can be seen as a more straightforward task, since there are somewhat similar views among different usability experts, but quality seems to have as many different definitions and ways of assessment as there are researchers handling the subject. Some useful discussion on assessing translation quality can be found in, for instance, Colina (2008, 2009), Schäffner (1998), House (1977, 1997) and Sharkas (2009).
Sonia Colina separates translation quality approaches into two categories: experimental and theoretical. Experimental approaches are described by Colina (2009: 237) as “ad hoc, anecdotal marking scales developed for the particular purposes of the organisation that created them, they suffer from limited transferability [...] due to the absence of theoretical and/or research foundations.” Alternatively, theoretical approaches “tend to focus on the user of the translation” and they “arise out of a theoretical framework or stated assumptions about the nature of translation” (ibid.). Colina argues that TQA research requires the following components:

- theoretical models and proposals that are verifiable and that pose clear research questions and hypotheses;
- theoretically and/or empirically based assessment tools, with clearly stated assumptions about theoretical or empirical foundations;
- evaluation proposals/tools that clearly state their purpose and limits;
- models/proposals that recognize many aspects of quality (componential)

(Colina, 2008: 103)

However, Colina (2009: 237) goes on to criticise some research-based functionalist approaches as follows: “[T]hey tend to cover only partial aspects of quality and they are often difficult to apply in professional or teaching contexts.” Regrettably, there have not been many studies on TQA from the users’ viewpoint. However, from the field of interpreting examples can be found that correlate with usability aspects. For instance, Kurz (2001) has studied what recipients of conference interpreting consider as good quality. She argues that assessing interpreting service quality should include the users and their expectations.
2.4.3 Quality Assessment Models

There has not been a single way to assess translation quality which would be universally applicable – nor in my opinion would one be likely to ever even exist. Here we shall take a look at some existing models of assessing translation quality and examine their relevance to a usability-centred evaluation.

When referring to past studies on translation quality assessment, Rodríguez Rodríguez (2007: 6) points out that “so far, most studies have only analysed the so-called mistakes of a translated text ... [which] has led to the study of other evaluative notions being ignored.”⁴ Here it must be noted that often the word ‘mistake’ does not appear as such in translation theory; for instance Nord (1997: 73) – among other scholars – uses the term translation error, which is seen not as a “mistake”, but as a “non-functional translation”. In fact, Nord proposes that “a particular expression or utterance is not inadequate in itself; it only becomes inadequate with regard to the communicative function it was supposed to achieve.” (ibid.) Inadequacy is seen as a quality assigned by an evaluator, not as a quality in itself. Therefore, translation errors should be seen as a larger part of a given translation, not just as mistakes, as Rodríguez Rodríguez points out.

One of the first names that comes up when looking at translation quality assessment is Juliane House, who has a long history in researching translation quality. Her book A Model for Translation Quality Assessment was first published in 1977. She has gone on to revise her model later, as can be seen in Translation Quality Assessment: A Model Revisited (1997). House’s model is known as the “functional pragmatic model”. House sees that translation quality assessment

---

⁴ Interestingly, Rodríguez Rodriguez’s focus is on the TQA of literary translations. She has aimed at creating a descriptive, contrastive model which includes the use of corpora as the means for the ST/TT contrastive analysis.
should focus more on a text-based approach instead of the target audience, and her main focus is on the relationship between the source text and target text and how they compare linguistically-situationally. However, in contrast to usability methods, House (1997: 159) sees the shift towards a target audience based approach as “misguided” and prefers using language experts as those who define translation quality.

House’s model also includes the aforementioned examination of translation errors. She divides errors into two groups: overtly erroneous errors and covertly erroneous errors (House, 1997: 45). The former group consists of text elements breaching the TL denotative meanings or language system. The latter in turn is seen as not succeeding in creating situational-functional matches in the TT. Similarly ‘covert errors’ as described by Vehmas-Lehto (1989: 2, 28–31) are ones that do not breach the TL language system, but differ from common use in the language; they “do not distort the message, but they hamper its communication” (ibid: 2).

When divided further, House's overtly erroneous errors consist of either breaching the language system or breaching “the norm of usage”, while covertly erroneous errors “demand a much more qualitative-descriptive, in-depth analysis” (House, 1997: 45). House does point out that the focus has often been too much on the overtly erroneous errors and that the weighting of errors in and between categories varies between each individual text.

Another quality assessment model which uses similar error identification can be found from the Copenhagen Business School (CBS) translation and revision process model and classification of errors, presented by Hansen (2008: 317–321). This model is developed especially for revision purposes when using the language pair Danish-German, although Hansen points out that the CBS classification of errors can be used "for all kinds of texts including the revision of literary works." (ibid.)
Much like House’s model, the CBS classification divides errors in two main groups and various subgroups. The main classification groups are 1) “errors in reflection to the affected units and levels of linguistic and stylistic description” and 2) “errors in relation to the cause ‘interference’ or ‘false cognates’”. The subgroups mentioned under the first main group are pragmatic errors, text-linguistic errors, semantic (lexical) errors, idiomatic errors, stylistic errors, morphological errors, syntactical errors and facts wrong. The subgroups presented under the second main group are lexical interference, syntactic interference, text-semantic interference and cultural interference. (Hansen, 2008: 320–322.)

However, this model presents only an equivalence-based approach and does not take into account the function of the target text. Hansen (2010: 385–386) does acknowledge the functionalist-based approach when describing different theoretical approaches to translation quality. She (ibid: 386) describes errors from a functionalist-based approach as “relative to the fulfilment of TT function and the receiver's expectations”, much like House’s covertly erroneous errors. Thus it could be said that in an equivalence based approach, the errors are identified from a language professional's perspective, while functionalism-based approaches are begging the question of reader response. It can be seen that these two approaches to translation errors overlap in many ways – it is mainly the focus of the evaluator that differs. To clarify this point, we could for instance consider a case where a language professional might notice some unidiomatic or ungrammatical use of language in a translation, but an actual reader would not be affected by it at all.

So, while House’s model is often quoted and used as a basis for other models of TQA, such as the aforementioned CBS model, it has also been criticised. For example, Colina notes that House’s model is based too much on the “notion of equivalence, often a vague and controversial term in translation studies” (Colina, 2009: 238). Equivalence-based approaches to
translation quality assessment are also criticised by Hönig (1998:23), since, in his view, they would only be applicable if it is assumed that the more equivalent a translation is, the better it is in quality. Of course, in favour of House’s criticism of moving towards an audience based approach, it must be mentioned that if the quality could be merely defined by the criteria of language experts, there would arguably be much more respect for the profession of translators and other language professionals.

Many others, including Colina and Hönig, have not been satisfied with previous translation quality assessment methods and have worked on developing these further. For instance, Colina has developed her own methods, the “functional-componential approach”. Based on her work regarding translation quality assessment, she has developed a TQA tool, which was originally created to be used when assessing the quality of healthcare education materials. The starting point for creating the TQA tool was a study regarding translated health education texts in the US, which identified translation quality as a problem; some of the analysed texts were in fact deemed almost unreadable without the ST. (Colina, 2008: 98.) It should be noted that similarly to Byrne’s examination of technical translation and usability (2006), Colina’s work with healthcare material also focuses on instructive texts.

Using Colina’s TQA tool requires both the ST and TT. The rating is carried out by reading the TT and ST and filling a form; the raters must be language professionals with native or near-native language skills in both the SL and TL. As raters, Colina has tested using bilinguals, professional translators and language teachers. The focus of the TQA tool is more on the translation (the product) itself instead of the translator and their actions.

The tool can nowadays be found, for instance, from the website of the Hablamos Juntos project (Spanish for ‘we speak together’), which aims to provide language services in health care,
especially in areas in the US with new and expanding Spanish-speaking populations. According to the downloadable manual from the project website:

The Toolkit is meant for translation requestors – individuals (or departments or organizations) responsible for initiating translations of health care text of all types whether they work directly with translators or through translation vendors.

(Hablamos Juntos, 2009.)

As can be seen from the above quote, the tool is something the recipients or customers of the translation product can use to assess the quality of the translations they require. Thus it does not take the user into consideration as such, but is more focused on the client, which correlates with Gouadec’s quality principles (2007).

Colina’s tool could be seen as an appropriate starting point for assessing translation quality, for it has been tested and piloted (Colina 2008, 2009). In addition, as can be seen from the Hablamos Juntos project, it is already in use. Also, while not using the term ‘usability’, there are similarities to be found between Colina’ TQA tool and usability evaluation. For more on the TQA tool, see Colina (2008, 2009).

We have also mentioned online crowdsourcing in the beginning of this chapter as a modern way of commissioning translations. In addition to using crowdsourcing and non-professional translators for translating texts, it has also been used as a method for evaluating translation quality to some extent. Chris Callison-Burch (2009) from the Computer and Information Sciences Department at the University of Pennsylvania has studied how Amazon’s Mechanical Turk crowdsourcing service can be used to evaluate machine translation quality. He found out that when the number of evaluators (“Turkers”) grew, their combined judgement was in close agreement with the evaluation gathered from expert computational linguists who work on machine translation. Callison-Burch suggests that this type of crowdsourcing is a cheap and
efficient way to evaluate machine translation quality, but does not comment on its use for human-produced translations. In addition, quality evaluation can be seen to be embedded into the crowdsourcing process used when translating Facebook, as presented by Mesipuu (2010). Mesipuu describes the translation process used as an “open community” crowdsourcing model (2010: 16), in which any member of the website can participate in the translating process. This results in various translations from different members for the same pieces of text. The quality evaluation aspects can be seen in the voting system, which the community members use to choose which translations they think are best (ibid: 20). Mesipuu adds that Facebook also does use in-house linguists to further evaluate and improve translations of certain major languages (ibid: 24–26).
3. Material & Methods

In this section we shall examine the material introduced in Chapter 1 more closely and describe how usability and quality assessment methods are used in this study.

3.1 The Guitar Handbook & Suuri kitarakirja

Ralph Denyer’s The Guitar Handbook, originally published in 1982, is an instructional book of guitar-related topics. It covers a wide range of different subjects, such as guitar playing, maintenance, famous guitarists and music theory. It has been well-received amongst readers. Its current average of customer reviews on Amazon.com is 4.7 out of 5 stars, where reviewers have commented it as “A Must Have” and “a great reference book” (Amazon.com, 2012). The book is described in the back cover as: “[A] handbook for players, as well as those interested in guitar building, repair and electronics […] the focus is still on the main issue – playing guitar […] the book is also great as a framework for self-study.” (Translation from Suuri kitarakirja by JS.)

The Guitar Handbook has been revised since its original publishing. The more recent English versions have copyright markings from 1982 and 1992. The articles in the ST have been updated in the later editions which can be seen, for instance, in the addition of new subchapters and the absence of some parts present in the first edition. Some changes which I noted when examining the different versions include more up-to-date information on recording technology and added or modified sections in the biographies of famous guitarists. When going through different Finnish editions of the book, it appears that the comments by Valkonen (2000) presented earlier are related to the first published editions of Suuri Kitarakirja. To give an example, the “freeze-dried” Frank Zappa translation mentioned in Chapter 1 cannot be found
as such in more recent editions of *Suuri Kitarakirja*. The translation used in this study is dated 1990, which suggests it has been translated from the pre-1992 first edition. Comparing this translation to the first edition of the ST, the layout and presentation seem to be similar and page numbers are identical. One might ask why the unrevised version is analysed in this study, since the assumed target audience would be more likely to have the revised version nowadays. The reason for the use of the original is because it would appear that most of Valkonen’s examples presented in Chapter 1 – which serve as an inspiration for this study – concern the unrevised translation.

When examining the forums of Finnish musicians’ internet forum muusikoiden.net, the book seems to receive much praise – it seems to be highly recommended as instructional material for guitar players –, but also ridicule for (at least parts of) its translation. However, when analysing the discussions, none of the forum members seem to be as harsh as Valkonen. To give an example, in *Suuri kitarakirja* the terms ‘pull off’ and ‘hammer on’ have been translated as ‘nauhanyhtäisy’ and ‘nauhaisku’ (p. 141). However, arguably most players in Finland would simply use the English terms to describe these techniques. The discussion forums mentioned above would indicate that the translated terms have mainly been a source of amusement for many members of the community, as can be seen from such comments as the following, written by user “Spiridon” on 17.10.2003: “No oilisin [sic] kyllä aivan äimän käkenä jos joku sanoisi että vedäpä parin [sic] nauhaiskut ja nyhtäisyt.”

In the case of *Suuri kitarakirja*, the translators do not appear to be as much professional translators as they are professionals of the guitar. To give an example, one of the translators,

Jyrki Manninen⁶, is better known as "Muddy" Manninen – a guitarist who has played with such bands as Wishbone Ash and many Finnish acts, including Hector (Wishbone Ash, 2008). This is an interesting point to consider, since these translations are made by people who, from an academically trained translator's point of view, could be considered as either non-professional translators or untrained translators (Susam-Saraeva & Pérez-González, 2012: 150–152).

Since this study focuses on finding out whether the book’s usability is hindered by its language, and not to assess the quality and usability of the whole book (which would be quite the workload), only an excerpt from the translation is chosen to represent the material. In addition, since the focus is on a summative evaluation (including adequacy and usability) of a final product, not linguistic equivalence, only the translated version is used for the tests.

The excerpt chosen for the testing is taken from a chapter concerning improvisation in guitar playing called “Yksiääninen soolotekniikka”, where different single-note solo playing techniques are presented. The test material contains pages 140–143. This section is chosen mainly because of its instructive nature, since instructive texts seem to fit usability testing well (as pointed out in Chapter 2.2). Using the selected material needs no previous knowledge of guitar playing techniques or musical theory; the user does not necessarily need to refer to previous information in the book. In addition, the material includes the aforementioned terms ‘nauhanyhtäisy’ and ‘nauhaisku’. The section is translated by Ilpo Saastamoinen.

---

⁶ His involvement in the translation has also been noted in a Wishbone Ash fan site: http://www.fubb.fi/sidelines/suuri_kitarakirja.htm
3.2 Methods

We have previously seen that empirical evaluation methods which include users seem to produce the most relevant and useful results for usability evaluation (Nielsen, 1997; Byrne 2006:180). Accordingly, many of the functionalist translation quality assessment methods discussed in Chapter 2 point towards reader-response in quality assessment. Here the focus is on the usability of the book, so usability evaluation methods are used. Since this study can only focus on the finished product, the testing is seen as a translation’s sumative usability evaluation. In addition, since the focus is on functional adequacy instead of textual equivalence, only the translation is used for the usability evaluation. The evaluation is performed in two major parts: (1) a heuristics-based expert evaluation and (2) usability testing. The methods for these evaluations are presented below. Their results will be presented and examined in Chapter 4, where we shall also discuss how well the methods worked and the success of the tests.

3.2.1 Expert evaluation

Four expert evaluators were chosen for the heuristic evaluation amongst members of the academic staff from the School of Humanities at the University of Eastern Finland. The evaluators were required to have experience with guitar playing and expertise in language. They were sent an excerpt from the book and an evaluation questionnaire, which is discussed in the following chapter. The experts’ answers to the questionnaire are used as the heuristic analysis of the material. Two of the expert evaluators have some previous experience with usability testing, but it should be noted that the evaluators here are not actual usability experts or novices (as suggested for heuristic evaluation), but primarily experts of language with a knowledge of guitar playing. Thus I have modified and applied heuristic principles to fit the
purpose of this study, using a questionnaire based on heuristics instead of having the experts perform the evaluation using a list of heuristics.

3.2.2 Heuristics

The heuristic principles used in the testing were modified from Nielsen’s (1995b) and Purho’s (2000) usability heuristics as well as Gouadec’s (2007: 6–8) translation quality principles, discussed in Chapter 2.3. They are largely based on Purho’s list of documentation heuristics, modified with Gouadec’s quality constraint principles to fit the material used. Hämäläinen’s (2008) study of applying Purho’s heuristics is also taken into consideration when applying the heuristics. From now on, I shall refer to the heuristics used in this evaluation as modified heuristics. These modified heuristics have been used to form a set of questions which correspond with applicable heuristics, as mentioned in the previous chapter. These questions are presented in more detail below, at the end of this chapter. Next we shall examine the principles used in combining the previous heuristics and principles into the set of new ones used in this study. Later I will clarify these by presenting a list of the modified heuristics and their definitions.

Gouadec’s (2007: 6–8) principles of a quality translation are incorporated partly into the modified heuristics. Since the translation is evaluated outside the translation process context, the principles of being “economically viable” and “compatible with the defence of the client’s or work provider’s interests” cannot be incorporated into the heuristics used here. However, Gouadec’s principles of being "accurate", "meaningful", "accessible" and "effective and ergonomic" are incorporated into the heuristics, as is the principle of being “compliant with any applicable constraint” to an extent, namely in terms which Gouadec describes as the “target communities’ linguistic and cultural standards and usages” (2007: 7).
The first modified heuristic is largely based on Purho’s “Match between documentation and the real world” heuristic, which is used to evaluate whether the words, phrases and concepts used in the material are familiar and whether the information is presented in a logical and natural way. This principle is also seen to include the evaluation of the language of the TT, which Valkonen (2000) describes as “impossible”. Purho’s first heuristic can be seen to correlate with Gouadec's principles of being "accessible" and “meaningful”. The similarities with accessibility are those of the information being presented in a logical way, the text being well-written\(^7\), and using language familiar to the user, which also correlates with the “meaningful” quality principle. In addition, being compliant with the target communities' linguistic and cultural standards and usages could also be seen to fit together with Purho’s first heuristic. Although Purho does not talk about compliance with any “applicable constraints” as Gouadec does, we can see how being compliant with the target culture’s rules and regulations and the translation’s physical limitations (in this case the book’s layout) are complying with matching the “real world”.

However, to keep the heuristics short and to the point, Purho’s first heuristic along Gouadec’s accessibility and compliance are used to form two separate heuristics, instead of trying to fit all under one principle. These two modified heuristics are called matching real world, which involves the compliance aspects and accessibility, which, in turn, includes the familiarity aspects and the text being well-written. However, since both heuristics are derived from the same source, I have chosen to examine these modified heuristics together in Chapter 4.1. Meanwhile, the matter of information being presented in a logical and natural way is moved

\(^7\) In Gouadec's (2007: 7) words “(preferably) well-written.”
under a different modified heuristic (information design), which is discussed later in this chapter.

The third modified heuristic is based on Gouadec’s principle of “accuracy”, which does not correlate as such with any specific one of Nielsen’s or Purho’s heuristics. This modified heuristic is named accordingly as accuracy. It requires the text to be factual and as free of technical and semantic errors as possible. However, as Gouadec (2007: 6) points out, a translation without any defects is hardly ever achieved and there might be errors in the original as well. For the purpose of this study, translation errors (as presented in Chapter 2.4.3) are not analysed here, but rather they are seen as a part of the previous heuristics of matching real world and accessibility.

The fourth modified heuristic is based on Purho’s third heuristic, “Purposeful documentation” alongside Gouadec’s principle of “effective and ergonomic”, to evaluate whether the intended function/use of the text is clear to the user. This principle also includes using appropriate media, which could play a significant role here – playing guitar has much to do with aural matters, which can be an issue for a printed book. Purho’s seventh heuristic, “Task orientation” which states that the documentation should focus on the task the user is using it for, is also included under this modified heuristic, considering its similarities with “purposeful documentation” and Gouadec’s related principle of “effective and ergonomic”. This modified heuristic is called purposeful and ergonomic.

The fifth modified heuristic is based on Purho’s “Support for different users”. Support is considered important, since the book is seen to be aimed at a large audience of guitar players with different levels of experience. This could also be seen to correlate with Gouadec’s “accessible” and “effective and ergonomic” principles. In addition, Purho’s eight heuristic,
“Troubleshooting” is not seen to be very relevant on its own, but there might, however, be parts in the text which could relate to support for possible problems, so it will be incorporated here as well. This modified heuristic is called **user support**.

The sixth modified heuristic is mainly based on Purho’s “Effective information design”, which evaluates whether the information is easily found and understood. It includes matters such as paragraph sizes and use of graphics. Some overlapping with the first principle can be seen here, when considering whether the sentences and paragraphs are easy to read. In addition, Purho’s sixth heuristic, “Support for various methods of searching information” is rather irrelevant in this study, taken that only an excerpt of the book is being used for the testing and evaluation. However, it is included here, grouped with the fourth and fifth heuristics and Gouadec’s principles of “accessible” and “effective and ergonomic”. I have also decided to incorporate the evaluation of “information being presented in a logical and natural way” under this heuristic instead of the already crowded first one, since presentation of information can be seen to correlate more with information design than matching the real world. This modified heuristic is called **information design**.

The remaining three of Purho’s heuristics are not included into these modified heuristics as such. However, Purho’s ninth heuristic, “Consistency and standards” is evaluated here by examining the consistency of the terminology. It is thus closely linked with the first heuristic, therefore it is included under the first modified heuristic, matching real world. Purho’s second heuristic, “Match between documentation and product” is not used in this test, since there is no actual “product” that the documentation is referring to which would involve written language. Similarly, the tenth heuristic, “Help on using documentation” is not seen to be of much use here, since the material only includes an excerpt of the book. If studying the book as a whole, this could be used in conjunction with the sixth modified heuristic, information
design. In addition, the tenth heuristic is described by Hämäläinen (2008: 84) as “rather irrelevant when evaluating small document sets.” He suggests combining this with Purho’s third heuristic, “Purposeful documentation”, since they cover similar aspects of the documentation. However, since this heuristic includes the possible updates on the documentation, it could be useful if different versions of the book were to be studied comparatively. In this study, Purho’s tenth heuristic is not, however, used.

As mentioned before, I shall clarify these modified heuristics below by presenting them as a list with definitions.

- **Matching 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 used effective.
Below are the heuristics-based questions that were sent to the expert evaluators, here translated from Finnish. They were formulated as open questions for the evaluators to answer in their own words. As the last part of the evaluation (not present in the list below) the evaluators were given the opportunity to give open comments on anything (or nothing) they wished.

- Are the words, phrases and concepts are familiar to the user? Are they are used consistently?
- Does the text deviate from the target language’s or community’s common usage?
- Are there factual or linguistic errors?
- Is the intended function clear to the reader? Does the given information focus on the purpose of use?
- Is the information suitable for users (players) with diverse levels of experience?
- Does the text provide support for possible problems which could come up while using it?
- Is the information presented in a logical and natural way?
- Is a printed book an appropriate medium for this purpose? Is the structure efficient (layout, use of graphics, paragraph sizes…)?
- Does the book seem appealing to use?
- Is the text well written?
- Did some specific problems arise concerning the use of the book?

These questions were chosen as a replacement for a list of predetermined heuristics to make the evaluation process less complex. As pointed out before, the evaluators were not usability experts, but had expertise in language and experience with playing the guitar. The questions were given to the experts in Finnish to be answered in Finnish, as was the evaluated material. The presentation of the questions does not necessarily follow the order in which I have previously presented the list of applied heuristics, for the questions are grouped according to topic – again to make the evaluation process more straightforward.
3.2.3 Usability testing

The usability testing is planned based on Rubin & Chisnell’s (2008) model presented in Chapter 2.3. The usability evaluation employs both observation and survey methods, with focus on the latter. The participants were found using personal contacts in Joensuu. They were required to have at least some level of guitar playing experience. To differentiate from the expert evaluation, language students and experts were ruled out. The four participants are all university students, aged 21–28. Three of them are male and one female; here all participants will later on be referred to with a generic ‘he’ pronoun, in order to prevent any identification. Their study majors include mathematics, physics and educational sciences. Three participants are completely self-taught, one participant has studied classical and electric guitar for a while at a Finnish music institute, but considers himself mainly self-taught. All have at least eight years of playing experience, however most comment that their level of playing activity has varied.

The usability testing consists of the participant using the book to practise playing techniques, followed by an interview. The tests were carried out one participant at a time on 11 and 12 March 2014, at a band rehearsal room in central Joensuu. The test sessions began with briefing the participant and giving them a written consent form. In addition, they were asked to answer written questions concerning their guitar playing history and possible experience with The Guitar Handbook/Suuri kitarakirja or similar learning materials. The participant was then given the same part of the book that was used in the expert evaluation and asked to use the material to practise the given techniques for up to twenty-five minutes – or less, if the participant deemed they were finished ahead of time. The participants were given the book itself, an electric guitar and an amplifier, a chair and a surface on which to place the book. They were also encouraged to vocalise their observations while using the material – however, the point
here was not to use TAPs as such (the vocalisations were not recorded), only to provide more data for the observation. The moderator was not to answer any questions regarding the information in the text – only to offer possible technical aid with the equipment used.

In the usability test, direct observation was used, without recording the participants’ performances. This decision was made to keep the test situation more comfortable for the participants – considering that some might not be comfortable having their performance recorded (especially those with less playing experience), and participants might incorrectly feel that their success and playing skills would be evaluated. It was also specified for the participants that their performance was not being evaluated, they were evaluating the text. A successful completion of the task is here defined as the participant understands the text and is able to learn the given techniques from it. If the participant is already familiar with the techniques, a successful completion is defined as the participant understands the text and sees it adequate for practising the given techniques.

The interviews were carried out immediately after the tests in a semi-structured manner; the questions were predetermined, but the possibility to follow up on interesting topics or ask new questions was kept open. The questions themselves are open questions, representing all three types presented by Byrne (2006: 189–190): factual, opinion and attitude. Unlike the other question types, which were presented orally, factual questions – namely background information – were asked on a written form alongside the written consent in the beginning of the test. The interviews were recorded with an audio recorder. The questions were formed to correspond in suitable parts with the expert evaluation, but some questions requiring language expert skills were left out. In addition, the questions for the users were made more subjective than those in the expert questionnaire.
The predetermined interview questions are as follows:

- Did you learn something new?
- How much did you rely on previous knowledge?
- Could the techniques be learned using only the book?
- Are the words, expressions and concepts familiar?
- Does the text seem effective to use, is it helpful?
- Does the book seem meaningful to use? Can you see yourself using it as practice material?
- Did some problems arise while using the book?
- Did you make mistakes and did the text give support for possible problems?
- Do you think a printed book is a good medium for practice material?
- What do you think of the structure – such as layout, use of pictures, paragraph sizes etc.?
- Is the text well-written in your opinion?
- Are there some improvements you would wish for?
4. Results and discussion

4.1 Results of expert evaluation

“The most glaring and funny parts of the [book’s] translation have been mercifully left out by the researcher.” Evaluator C.

In this section, I shall examine the expert evaluators’ answers according to the corresponding heuristics. All quotes have been translated from Finnish by JS. When needed, the evaluators will be referred to as evaluator A, B or C.

Three of the four faculty members chosen for the expert evaluation provided their answers within a month of sending the evaluation form and text excerpts. The fourth evaluator had to withdraw due to time-related issues. All evaluators stated they had over 20 years of experience with guitars with varying levels of playing activity. All evaluators had previous experience with some version (Finnish or English) of the book – this is also evident from evaluator C’s comment above. Two evaluators mentioned that they had some experience in usability testing. Evaluator B stated he owns the English version of *The Guitar Handbook*.

4.1.1 Matching real world & Accessibility

The first modified heuristic requires compliance with the translation’s physical limitations and the target community’s rules, regulations and linguistic and cultural standards. There are similar points of interest between this heuristic and the following heuristic, *accessibility*, which requires that the overall language is familiar, including words, phrases and concepts, which should be used consistently, and that the text should be well written. In addition, these heuristics are arguably the most interesting concerning this study, since analysing accessibility here seems to present some suggestions to whether Valkonen’s critique of the language making
the understanding of the translation difficult is justified. Thus the two heuristics are examined here together.

The evaluators noticed some deviation from the common use in the target language and community. Evaluator A comments that there are some “awkward” and “archaic” features in the word choices and syntax, giving examples such as the word order (placement of the verb ‘ovat’) in this sentence from page 142: ”Yleensä ovat nousevat liu’ut helpompia kuin laskevat.” Evaluator B comments that the text would be more in line with Finnish customs if it were more straight-to-the-point and that referring to the reader in second person is not as common in Finland as it is in anglophone cultures. Evaluator C states that the overall text is not very good Finnish and that the excerpt used in the evaluation has “an impression of a word-to-word translation”.

All evaluators agreed that the terminology used in the text would be familiar to the user – at least the original English terms, which in the translation are presented alongside the Finnish ones. However, some of the Finnish equivalents were considered strange, even amusing. These include “nauhaisku” and “nauhanyhtäisy”, which received comments from all three evaluators. The evaluators agreed that the terminology would be more familiar if only the English terms were used instead of the translations. Evaluator C considers the translated terminology as “sympathetic” and an “innovative attempt”, but points out that some of the terms used in the translation are not (and have not become) established terminology. The terminology is deemed to be overall consistent, however, evaluator C notes that when referring to string bending, the translated text occasionally uses “taivuttaminen” as opposed to “venytäminen”, which the evaluator sees as the more frequent and common term.
Evaluator C comments on various problems with unnecessarily difficult language. Some examples include noun-heavy expressions such as ”[k]ielen venyttäminen suoritetaan”, “tapa välttää kielen epävireeseen menoa” and ”kieltä on mahdollisuus venyttää.” Similarly, evaluator A comments that some translated sentences seem illogical, such as ”[k]eskisormesi ei pitäisi osoittautua liian hankalaksi” (p.143). Evaluator C also notices source language interference in the translation, pointing out these examples from page 142: “Sormen puristus on nousevien niin kuin laskevienkin liukusävelten avain” and “kevytkosketuksisempin kielin.” He suggests that the translator should have focused more on translating the message instead of the words.

Interestingly, when considering the text from a broader angle, evaluator B states the text overall “needs work on the language, but this does not affect understanding the subject.” Similarly evaluator A believes that the written language “is not at all terrific, but not so bad that it would affect understanding.” In his opinion, “the text manages to present the information, but it is in no way a great reading experience.” Here, evaluator C is, to an extent, in line with these comments and states that the text is well-written, “excluding [my] previous comments on translation and unnecessarily difficult language.” He describes the text’s tone as “motivating and inspiring, despite its awkwardness.” Evaluator B confesses to having taken a look at the original English text, whose language he describes as “fluent and natural”.

4.1.2 Accuracy

The accuracy heuristic requires the text to be as factual and as free of technical and semantic errors as possible. It should be pointed out here that while covert translation errors (as presented in Chapter 2.4.3) could be seen to fit under this heuristic, I have decided to analyse related issues such as difficult sentence structures under the accessibility heuristic.
The evaluators agree that overall the text is factually accurate. Evaluator A does point out some parts of the text that could be open to interpretation. This includes a part from page 141, where the text claims that pull-offs and hammer-ons are limited to a space of four frets (this relates to how far fingers can reach, translation by JS) – whereas evaluator A states that this is not the case, especially near the 15th fret, where the space between the frets is much narrower. The text also suggests that the most effective chord slides are on fourth, fifth and octave intervals, which evaluator A sees as a subjective opinion presented as a fact. I wish to comment, however, that these are the opinions of the original author, not the translator.

Evaluators B and C comment that the text itself does not contain almost any factual errors, but there are problems in spelling and grammar (as seen when examining the accessibility heuristic in the previous chapter). Evaluator B states there are misplaced or missing commas here and there (but does not give examples) and points out a few typos on page 140, such as: “[…] soolossa pitäisi tapahtu jotain.” Evaluator C describes the text as “thoroughly compiled” and praises its factuality, but criticizes the language of the translation as “quite awkward gobbledygook.” In addition, evaluator C notices a part where the text might give an “unrealistic” or “inconceivable” impression between a picture and the written instructions (p. 143).

4.1.3 Purposeful and ergonomic

This heuristic includes three aspects: 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.

8 “[… ]aika kömpelöä kapulakieltä.”
All evaluators agree that the function of the text is very clear to the reader and the information is mainly effectively focused and relevant. However, evaluator B does notice that the presentation of information could be more straight-to-the-point to be more compliant with Finnish cultural standards regarding instructive texts. In addition, evaluator C points out that the description of pull-offs includes some unnecessary information (performing the pull-off in a 45 degree angle, p. 141) that can be more a distraction than help for a player learning the technique.

The book format is overall seen by the evaluators as a good medium for the purpose. In evaluator C’s words, “[a book’s] use does not require external devices” and it “does not create electric hum while playing.” Evaluator B also praises the book as an interface, which is easy to carry around and does not require electricity. However, all evaluators mention videos as a more illustrative option than pictures in a book – but as evaluator C points out: “It's unbelievable that there were no YouTube videos in the past.” Similarly, evaluator B suggests that using some type of audio examples would be a good addition, since “after all, music is primarily sound.” Evaluator A comments how “a book and video clips would support and complement each other as self-study material.”

### 4.1.4 User support

This heuristic states that the information should be suitable for users with different levels of experience; the text should also provide support for possible problems that might arise while using it.

The evaluators agree that the information on the chosen pages is most suitable for beginners. As evaluator A suggests: “Not many professional players would get new stimuli from these sections.” Evaluator B does mention that page 140, with comments from professional players,
could be of interest for more advanced players as well – pointing out Richie Blackmore’s suggestion that guitarists should listen to saxophone solos. In addition, evaluators B and C agree that the book as a whole contains useful information for players with various levels of experience.

Evaluators A and B see that some attention has been paid to possible problem situations. Evaluator A mentions as an example the description of how to avoid possible unwanted sounds when bending strings. Evaluator B, however, suggests that a reader familiar with the subject might not even recognise possible problems that a less experienced player might face when practising the techniques with the book.

Evaluators A and B think the pictures might not be clear enough at first, however, evaluator B sees that the text explains the pictures and that the pictures provide additional clarification. Evaluator A would have preferred pictures taken from a player’s point-of-view instead of only pictures taken from the front, which might not be the best option in his opinion.

4.1.5 Information design

This final heuristic concerns information, which should be easily found and understood, presented in a logical and natural way, including effective use of graphics and paragraph sizes.

The information flow is deemed logical and natural by all evaluators. Evaluator A sees that the information is presented in a clear and logical fashion, but criticises the layout – the information is divided into different sections which are presented inside separate boxes; this creates a fragmented impression in evaluator A’s opinion. Evaluator C shares this view and
suggests that using separate divided sections to highlight information loses its effectiveness when the body text is presented inside these boxes – he also suggests that the body text and text boxes are not aligned properly in the chosen excerpt. Evaluator C calls the layout and design as “not especially interesting or vibrant” when compared to modern school books. However, it should be noted that while these issues with the layout can be seen as usability problems, the translation itself could still be seen to comply with the physical limitations presented by the layout, since they are the same in the ST and TT.

4.1.6 Summary of expert evaluation

To sum up, the expert evaluators find the text to be understandable, but criticise the language. Most problems relate to language features, such as sentence structures and word choices. Some of the translated terms, especially ‘nyhtäisy’, are seen to be strange or even amusing. Some parts of the text are considered open to interpretation and some subjective opinions are presented as facts. The expert evaluators see the chosen text excerpt to be mainly suitable for beginners. The layout presents some problems, the pictures are not very clear, but the information flow is otherwise seen to be logical and natural. It should be taken into consideration that the language problems could have been remedied in the translation process, while aspects such as layout, factuality and medium are features of the ST, and as such difficult for the translator to influence.
4.2 Results of usability testing

In this section, I shall examine the data gathered in the usability tests. All quotes have been translated from Finnish by JS. Here the participants will be referred to as participants One, Two, Three and Four, using ‘he’ as the gender-neutral pronoun for clarity and anonymity. Only one of the participants has any significant previous experience with Suuri kitarakirja – he owns a copy, but has not read it in a while. Another participant has skimmed the book a few times at a friend’s house. All have used some similar self learning material, such as Blues Station by Harri Louhensuo, which came up in two of the interviews. However, most participants agree that online guitar tablatures and playing videos have been a more common means of self-education than printed books. The interviews provided the most useful data, but the direct observation provided some interesting observations, too, which will be discussed here first before proceeding to the interviews.

4.2.1 Results of direct observation

The way the participants went through the text varied quite a lot. Two of the participants proceeded linearly through the text and practised the techniques as they came along, one participant read the whole text through before proceeding to apply the techniques, and one skimmed the text first and proceeded to practise the techniques in a non-linear order. Similarly, two of the participants used the techniques only as they were presented in the text, i.e. playing nothing but what was described in the text, while the other two used them as parts of improvised playing. One participant was playing something on the guitar much of the time while reading the text. The book seemed to work well with playing, placed on a surface in front

---

9 Tablature is a form of musical notation that describes instrument fingering instead of musical pitch. It is most commonly used with fretted string instruments.
of the user: there was no problem keeping it open or turning the pages and it did not interfere with holding the guitar. However, its small print produced some problems, which will be discussed later in this chapter. The time each participant spent on the text varied between 10 to 25 minutes.

All participants were familiar with the techniques, so not many errors were made. Only the part describing the vibrato technique presented some problems. The vibrato technique is presented as being possible to perform either horizontally (along the guitar neck) or vertically (by bending the string sideways). The former technique was unfamiliar to most participants, and while all managed to perform it vertically, the horizontal vibrato presented some problems. Participant Three commented that while the text describes the horizontal vibrato as the more common one, he has never used it himself and that it would likely suit an acoustic guitar better than an electric one.

The participants were mainly quiet while going through the text, although they were encouraged to speak up if something caught their attention. Some of the vocalisations included unfamiliar information in the text or parts that were considered amusing. Participant Three said he had never encountered the term “nauhanyhtäisy”. Similarly, participant Four had never played the hammer-ons and pull-offs in one of the ways presented in the text – as a descending or ascending four-fret chromatic progression – and said “this is turning quite jazzy.” While reading, he also criticised the large amount of text being packed into a small space.
4.2.2 Results of interviews

The successful completion of the test was here defined as either:

- the participant understands the text and is able to learn the given techniques, or
- the participant understands the text and sees it adequate for practising the given techniques.

Because all participants were familiar with the techniques, the latter definition was applied. Since evaluating this definition requires data gathered during the interviews, we shall now move on from the direct observations to examine the results of the interviews. Accordingly, all participants completed the test successfully. Most of the participants did not learn anything new, but the text was helpful for at least participants One and Two. Participant One said he had a chance to brush up on skills they had not used in a while. Participant Two commented that the book gave him some specific details that he was not familiar with before – as he commented “fine-tuning for the techniques”\textsuperscript{10}. Participants Three and Four stated that while they did not learn anything new, the book could be used to learn the techniques. Still, some usability problems did occur, which we will return to shortly.

Interestingly, none of the participants had negative comments about the language itself and all describe the text as well-written. Participant One saw the tone as inspiring and motivating and commented that the detailed explanations help understand why the techniques are done the way they are – in their words “the author wishes to help, not to boast.” Similarly, participant Four said he likes the tone of the writing and that the language is overall good Finnish. Participants Two and Three said that nothing in the text caught their attention and that the text is on the whole good. Participants One and Three thought the educational style of the text was fit-for-purpose.

\textsuperscript{10} Originally: ”Hifistelyä siihen tekniikkaan.”
However, the sheer amount of text presented some problems. Participant One stated, “a secondary school student interested in the subject [...] would be a little struck with despair, as this was supposed to be fun, but it turned out to be school study.” Participant Four commented that there is too much content on the four pages, which might make it difficult to focus, and suggested that such subjects are much easier learned by seeing instead of reading. In their words “this would give a discouraging view [for] someone who has never played guitar before, a terrible pile [of text].” Participant Four also criticised that (assumedly) all the information on these subjects was compressed into one single section of the book. In contrast, participant Two commented that using previous knowledge on the subject helped the reading process, giving the reader the possibility to skip parts of the text, since the flow of information is logical. Similarly, participant Four did agree that the book would be better suited for someone already familiar with the subject – they suggested a guitar teacher searching for ideas could benefit much from going through the text.

Participants One and Two did comment that the large amount of text in small print made the book somewhat difficult to use, since while holding a guitar, the reader has to reach near the text to be able to read – participant One suggested this could be helped by raising the book eye-level with a music stand. Participant Three saw the layout choices and length of the chapters as “justified”, and considered the boxes in which separate parts were presented as “clear”, as did participant Two. On the other hand, participants One and Four found the layout “heavy”.

There were also some problems seen in the placement of the pictures in relation to the text. Participant Four commented that the pictures are “scattered around” and that their context can be weak, which affects concentration. Participant One stated it took him a while to find out which part of the text the pictures related to, but did not see this as particularly problematic.
Participant Two commented that in some pictures it is difficult to make out what is happening or which part of the text the picture is referring to. However, three of the participants liked the pictures for the most part and found them quite helpful. In addition, participants Three and Four said they were not familiar with some of the symbols and illustrations used alongside the pictures and thought that these might be explained in some previous section of the book – which, in fact, is not the case.

The terminology is generally familiar to the participants – especially the English terms. However, “nyhtäisy” is commented by two participants as an unfamiliar translation for pull-off. Both find it quite amusing, in participant Four’s words it is “quite a fun translation.” Participant Three says such a term was not used when he was studying guitar at the institute. Participant Two, on the other hand, suggests he is familiar with the term.

None of the participants suggest much improvements for the text as such. Three of them did comment that some audio-visual material to accompany the text would be useful. Participant One suggests that the part of the text describing the vibrato technique should be moved before the part on sliding – for vibrato is referred to in the slide-section. Participants Two and Four would have preferred tablature notation instead of the unfamiliar illustrations. One and Two suggested a larger print would help reading the book while holding a guitar. Participant Four would have preferred the information to be spread wider throughout the pages and to “concentrate on the level at hand, keeping the jazz level in the jazz sections and basic level on basic sections.”

11 Participant Four was more critical of the pictures, suggesting that “the pictures are shit, you can’t see what is going on.”
4.2.3 Summary of usability testing

All in all, the participants find the text to be mostly helpful, but there are some problems in its usability. Most of the problems concern the layout; a large amount of text printed in a small font can make the reading difficult, the extensive quantity of information can be troublesome to follow and the pictures are not necessarily logically placed within the text. Interestingly, the majority of the problems were not related to the language of the translation. The users found the text to be well written and the language to be overall good. However, the term ‘nyhtäisy’ was seen to be strange or amusing, similar to the expert evaluation. In addition, most participants did comment that a book would be less familiar learning material than online tablatures and instructional videos, which could link to the age of the test participants compared to the year of publication of Suuri kitarakirja and the somewhat older generation of guitar players represented by the expert evaluators.
4.3 Discussion

This section is divided into two parts: discussion of the primary and secondary purpose of this study. In the first part we shall look at the results of the two tests in light of the previously discussed usability evaluation. In the second part, the results are compared to functionalistic TQA, the tests themselves are discussed and the application of the methods and their success is examined.

4.3.1 Discussion of the primary purpose

In this section, the primary purpose of the study, analysing *Suuri kitarakirja* with usability methods, will be addressed. We shall begin with having an overall look at the results of the expert evaluation and usability testing. The usability of the text will then be analysed more closely according to Byrne’s and Nielsen’s definitions. We shall also examine some differences between the user participants’ and expert evaluators’ opinions.

The usability problems discovered in the two tests were not as devastating as Valkonen’s critique would suggest, although some issues were found. The expert evaluation did not include a severity rating system (introduced in Chapter 2.3.3), but the evaluators’ comments would suggest that none of the usability problems found were “major” or “catastrophic”, to use Nielsen’s ratings. In fact, most of the larger problems found in the expert evaluation are language-related, while the language as such did not seem to present problems in the usability testing, where it was deemed by the users to be overall good. On the other hand, the participant users did find some layout features problematic: the large amount of text in small print presents some problems, and the pictures can be hard to understand because of their positioning and small print. These layout problems also came up briefly in the expert evaluations, although they were much more prominent in the users’ comments. Interestingly, the test situation might
assumedly draw attention away from the language features and bring more focus to other aspects of the text. This is an interesting subject which would benefit from being studied further.

Interestingly, it would appear that the translation is not as bad as Valkonen suggests – at least judging by these results using the four-page excerpt. While Valkonen's description of the language is “impossible” and “all but unreadable”, the evaluators' opinions are less harsh. They see the language as “awkward” and “not at all terrific”, but still consider that it presents the information without overly disturbing the message. In addition, the usability testing also suggests that the information is presented well and the language does not interfere with using the book.

To evaluate the usability of the book more closely, we shall return to Nielsen’s and Byrne’s definitions of usability. Nielsen’s (2012) definition consists of five quality components – satisfaction, errors, learnability, efficiency and memorability – that define a product's usability. Byrne’s definition states, “[w]hen 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). There are many similarities between these definitions, but I wish to focus here on Byrne’s definition, since it is arguably easier to apply to texts. In addition, Nielsen’s memorability could not be evaluated here, since only one session was conducted with each user.

To apply Byrne’s definition, it shall be broken down into fragments and analysed accordingly. The first part of the definition measures the extent to which the reader can read the text. Here no major problems are detected, although some problems are found in the text’s layout. The
small font and large amount of text could hinder reading in a situation where someone is using the book to practise playing the guitar, since the reader would have to be close to the text to read. The evaluators’ comments would suggest that the language of the translation might interfere with reading, however, this was not apparent in the usability testing. Some of the users and evaluators did have trouble combining pictures with corresponding parts of the text, which also hinders reading to some extent.

The second part of Byrne’s definition measures the extent to which readers can understand the contents of the text – this corresponds to Nielsen’s learnability. I will include understanding the pictures and illustrations here as well. Again, the language does not seem to interfere with the users understanding the contents. This is suggested by the expert evaluation as well as the usability tests. The only noticeable problems with understanding seem to be in relation to the pictures and their placement within the text.

Byrne’s definition continues with the extent to which the reader can perform the task required by the text quickly and accurately – corresponding with Nielsen’s efficiency and errors. All users were able to successfully complete the task according to the predetermined definition. The time taken for them to go through the text and be satisfied with their results varied from 10 to 25 minutes. However here it must be considered that all users were familiar with the subject – as participant Four pointed out, these are techniques that many players spend years perfecting and here the information is fit into four pages. The only notable problem with performing accurately seemed to be the horizontal vibrato technique, which was unfamiliar for most participants, since they were more used to performing it vertically. Regardless, the users’ performance could overall be described as quick and accurate.
The final part of Byrne’s definition is the extent to which the readers find the experience difficult or easy. I will include Nielsen’s quality component “satisfaction” here, since apparently it is not present as such in Byrne’s definition. Interestingly, most users had no problems with using the text themselves, but deemed it might be more problematic for younger or less experienced players. Similarly, both tests suggested that using nothing but a book to learn the techniques would not be an ideal situation for an inexperienced player – instead, some audio-visual material or the help of a more experienced player were proposed as an alternative. Indeed, both tests did suggest that a book might be somewhat outdated as learning material for guitar playing. Here the evaluators’ comments on the language being difficult or awkward might affect the experience, although it did not affect the participants of the usability test. The participants’ comments on the layout being “heavy” can also affect the experience.

It could be argued that in this case, most of the problems in usability had to do with the experience, or satisfaction – at least on the basis of the results of the usability tests. This can be seen particularly in the participant users’ comments presented in Chapter 4.2. Then again, in view of the results of the expert evaluation, most problems concern the “accessible” heuristic. This is, indeed, in line with Valkonen’s criticism. However, the degree of the problems is debatably much lower in these results – yet it must be taken into consideration that this study included only a small sample of the translation. Most usability problems found in the expert evaluation were focused on language, mainly within the “accessible” and “matching real world” heuristics, while the other heuristics could be deemed less problematic. Similarly, possible problems with compliance with the target community’s standards arose in the usability testing, where the participants commented they were unfamiliar with some illustrations used in the text and would have wished for tablature notation and/or audio-visual material to accompany the text.
In terms of the “accuracy” and “purposeful and ergonomic” heuristics, the usability of the text can be considered good, with only minor complaints from the experts. Similarly, in light of the “information design” heuristic the usability was deemed as good, but not especially great and the “user support” was okay usability-wise, although the experts found the text mainly suitable for beginners. Interestingly, while the experts seemed to share the view that the text would be mostly suitable for beginners, some of the users’ comments would indicate that the text might be off-putting for a younger, less experienced player, and would be more beneficial for readers already familiar with the subject.

However, it must be stressed that most of the usability problems that were found in the tests were not necessarily connected to the translation. Indeed, problems which could be caused or fixed in the translation process were mainly found in the expert evaluation, not in the usability testing. Similarly, some of the problems, such as preferring audio-visual material and tablatures alongside the text, could be related to developments in technology and changes in guitar players’ self-learning methods. Arguably the expert evaluators could be seen to be closer to the book’s target audience than the younger test users, considering that the book was first published in 1982.

4.3.2 Discussion of secondary purpose, methods and success of the tests

This section addresses the secondary aims of this study: the application of usability evaluation methods and comparing them with theoretical TQA models. First, we shall discuss the success of the test, examining the methods and their application. Second, the results are viewed in the light of functionalist translation theory and translation quality assessment.
The expert evaluation was arranged in a way that differs somewhat from the ideal situation suggested by Nielsen (1995a) and Suojanen et al. (2012: 101). While the number of evaluators is sufficient (three), further evaluation had to be discarded due to available time and resources. Some features that had to be left out include the severity rating of usability problems and having the experts discuss their findings together after individual evaluations. How much more useful data these further evaluations would have elicited can be debatable, but for the scope of this study, I believe that the expert evaluation methods here are adequate. Indeed, considering that the secondary purpose of this study was to test usability methods, I believe the results were sufficient and the study provided interesting data of applying usability-based expert evaluation. In addition, the required expertise of the evaluators used here could be questioned in the light of the definition given by Suojanen et al. (ibid.), which states the evaluators should be usability experts, novices or experts with knowledge on both usability and the evaluated product. Here, only two of the three evaluators have experience with usability, but the main criterion is their experience with language and the “product”, or guitar playing. Personally, I believe the experts fit the purposes well and the data they produced was useful.

However, the choice of building a modified set of heuristics proved to be somewhat more problematic. Whether I should have kept to one of the pre-existing heuristic lists can be debated. Nevertheless, I do not feel that using Purho’s or Nielsen’s heuristics as such would have been as beneficial, considering for instance Hämäläinen’s (2008) study on Purho’s heuristics. In addition, since I did not use usability experts in the evaluation, I decided to form the heuristic evaluation into a set of questions which address the heuristics. This, in my opinion, worked out rather well, since the evaluators had no problems answering the questions and their answers corresponded well with the chosen set of heuristics. However, the heuristics themselves could benefit from more refinement. Now a deal of overlapping can be detected between the heuristics, and some of the heuristics are arguably too extensive. This can be seen,
for instance, when examining problems with syntax, e.g. whether overly difficult sentence structures should be examined under “errors” or “accessibility”. Indeed, the “errors” heuristic could benefit from applying the overt/covert error aspects of translation errors and by specifying the differences between ‘errors’ and ‘the text being well written and familiar to the reader’. I believe the heuristics used here could be broken down into a larger set of more specific heuristics. In addition, since this list was created specifically with this study in mind, its wider applicability can be questioned. Regardless, the data gathered by using the modified heuristics and the questions addressing them proved to produce useful material for the study.

The usability testing meets Nielsen’s suggestion of having 3–5 participants per test. The group can be considered (at least roughly) a specified target group: all participants are university students in their 20s with more than 8 years of guitar playing experience, with varying levels of activity. This, and the fact that none of them are students or experts of language, distinguishes the user group from the expert evaluation group who are language experts with over 20 years of guitar playing experience. This distinction between the backgrounds of the participant users and the expert evaluators provides an interesting variation, which could have an effect on the data gathered from each group. However, a different user group could have provided different results, as suggested by the participants when considering younger players unfamiliar with the techniques in the excerpt. Now the user experience is largely defined by one distinct target group.

Interestingly, performing the usability testing at a rehearsal room raises some questions. It can be seen as a realistic place where guitar players would practise, but as a setting for usability testing there were some problems. Two of the test sessions were held late in the afternoon, and as a result, they were somewhat affected by the noise coming from the adjacent rehearsal rooms, when bands came to play. Furthermore, the positioning of the user, book and
instrument had to be improvised – for the lack of table, an overturned guitar amplifier was used as a makeshift table in front of the chair the user was sitting on. Nevertheless, the users seemed to enjoy the test situation and environment and were happy to participate.

The survey methods (interviews) used in the usability tests proved to offer more suitable data for the study than the direct observation – the survey methods were emphasised, after all. The application of more observation methods could have provided some interesting data, however, for instance video recording all the users’ actions when practising the techniques would have provided much data for analysis, but might have also affected the users’ performance. The interviews provided a great deal of interesting data, however the questions themselves would have profited from refinement. Fortunately, the interviews were kept semi-structured, so interesting discussion topics were able to be fit in the interviews when they came up. In addition, while three of the participants were quite content with the book, the fourth was more critical and less satisfied with the experience. This could be seen to correlate with Rubin & Chisnell’s (2008: 72–73) suggestion that using more participants can reveal problems not found otherwise. However, obviously such a small number of participants with individual opinions cannot be considered statistically significant, although Nielsen (2012) maintains that for finding usability problems, no more than five participants at a time are required to acquire “maximum benefit-cost ratio.”

Now we will take a look at how the results would look in view of theoretical TQA models. As seen when discussing TQA models in Chapter 2.4.3, in an equivalence-based approach, the errors are identified from a language professional’s perspective, while adequacy-based functionalism includes reader response. In this study, by employing usability methods, we have left out the problematic and, in my opinion, somewhat outdated equivalence-based approach; however, some similarities can be seen between an equivalence-based approach and the expert
evaluation (language professionals’ perspective), compared to functionalistic models and the usability testing (reader response) applied here. Similarly, although equivalence was not directly assessed in the analysis, some of the experts did comment on language features which would be part of an equivalence-based approach, such as interference of the source language (see e.g. comments by evaluator C in Chapter 4.1.1). It is worth considering that while focusing on a summative evaluation of the product, we could not address the client who commissioned the translation as such. The client is arguably one of the most important factors in functionalist translation theory, such as Gouadec’s (2007: 6–8) translation quality principles. Attending to the client would be an interesting subject to add to future studies of translations’ usability evaluation.

From a skopos-based viewpoint, it could be argued that the text fulfils its purpose – if the purpose of the book is seen as a handbook for guitarists and a means of self-study, as pointed out in the back cover (see Chapter 3.1). Similarly, the translation of the excerpt used could be seen as adequate, since the usability testing suggests it fulfils the communicative act of presenting applicable self-practice material for guitar players. Arguably the evaluators do not appear to consider the translation inadequate – if inadequacy is seen as a quality assigned by an evaluator –, since they seem to agree that the translation is understandable, while not ideal, language. In the case of translating Suuri kitarakirja, communicating the guitar-related contents could be seen as a more important task than achieving a high quality of language. This is reflected in hiring professional guitarists instead of professional translators to translate the text. In the expert evaluation, the translation was, indeed, deemed to be factually accurate. Personally, I do not believe that finding one or more professional translators with expertise in guitars would have been an impossible task for the client, which would have solved both language quality and factuality issues – conversely, a language professional could have checked
the guitar experts’ translation. However, again I must note that these assumptions are based on the expert evaluation and usability testing, since the client’s role has not been addressed.

When considering errors, which, as seen in Chapter 2.4.3, have been an important part of previous models of TQA, we cannot see many cases of overt errors. This can be mainly due to not using an equivalence-based approach, however, for instance the term “nauhanyhtäisy” could be taken as an overt error based on House’s (1997: 45) description of breaching “the norm of usage.” Instead, a more careful analysis could present many covert errors in the text, based on the expert evaluation – keeping in mind here how Vehmas-Lehto (1989: 2) describes covert errors: “[they] do not distort the message, but they hamper its communication.” However, since analysing errors was not in the centre of this study, we shall not pursue this notion further here. Instead, I wish to comment that usability evaluation would seem to present an applicable angle to tackle TQA – indeed, it includes such aspects, for instance, extralingual factors (layout, appropriate medium, user support etc.), which cannot be assessed in equivalence-based models.

Another question to consider is how well the usability of the text can be evaluated by using these methods and how can it be compared to TQA, as discussed above. Now, since the problems identified by the expert evaluators were not rated for severity, a statistical comparison cannot be made. The results have to be analysed by the researcher according to the comments made by the evaluators, and usability problems cannot thus be objectively arranged. Since the focus is on usability, the question of traditional TQA cannot be directly addressed. Indeed, much of the usability problems found in the usability testing are arguably not related to the translation as such, which is quite noteworthy, considering that traditional TQA would not have even addressed the types of problems found in this study – examining usability does give a somewhat unique perspective on the functional quality of translations. Of course, if we
would have wished to focus more on traditional TQ, the heuristics and the interview questions could have been modified to address translation quality, yet, I did not wish to lead the users and evaluators to focus too much on the translation itself, unless it happened naturally – as it occurred in the evaluators' comments presented in Chapter 4.1. Interestingly, the participant users did not comment much on the translation, they noticed the translation aspects mainly in the terminology.

In the light of this study, I do consider that TQA methods could be developed and improved using usability evaluation methods and that usability and usability evaluation methods could be beneficially incorporated into functionalist Translation Studies.
5. Conclusion

The primary purpose of this study is to evaluate an excerpt of *Suuri kitarakirja* by using usability methods. The results of the expert evaluation and usability testing would suggest that while the language of the translation is not very good, it does not affect the reader understanding the information. Most of the book’s usability problems found in the expert evaluation concern language matters, but these did not come up in the usability testing. Instead, the users do not have complaints over the language and consider it good, except for the term “nauhanyhtäisy” which was unfamiliar and amusing for three of the four participants. This is somewhat contrary to Valkonen’s critique, which suggests that the language of the translation makes it nearly impossible for the reader to understand. The problems found in the usability testing seem to focus on the large amount of information being packed into a small space, written in a small font. The secondary purpose of the study was to test usability methods. The usability methods seemed to suit the evaluation the book well, however in this study they arguably did not directly assess the translation itself.

There are many different ways in which this type of usability-based translation evaluation could be studied further. For instance, the heuristics used in this study could be improved and made more widely applicable for similar texts or even for various text types. Similarly, it would be interesting to combine the heuristic evaluation methods with functionalist TQA methods, such as Colina’s (2008, 2009), in order to make assessment methods focused on the translation itself. A pre-existing TQA method could also be used side by side with a heuristic evaluation and/or usability testing, applying them to the same translation and seeing how similar or different results they produce. This type of comparative study could also be a basis for usability-based TQA models. An interesting point to consider is also whether usability methods could be broadened from single texts to focus on large masses of text (corpora).
Of course, *The Guitar Handbook* could also be studied further, using revised editions and larger parts of the text than the four-page excerpt used in this study. For instance, usability aspects of different editions of the translation or between sections by different translators could be evaluated. In addition, a ST/TT comparison could also be interesting, and its expert evaluation results could be compared to those gathered using only the TT. However, if the book were to be studied further, I would prefer to see more focus on improving evaluation methods instead of merely analysing quality aspects of the translation. I do agree with Valkonen that the book could be used as learning material – it might be used to demonstrate how translation competence grows in translator training, for instance, by comparing the book’s translation to students’ translations of the same ST. *The Guitar Handbook* would present a difficult special-field text, whose translation would seem more focused on factual accuracy instead of fluency in the target language.

To conclude, I wish to point out again that, for instance, House (1997: 159) does not give much value to the audience-based approach, but prefers using language experts as those who define translation quality. While this would be an ideal situation, regrettably I do not see it such an easily applicable practice in modern translation industry. Instead, I would personally wish to see more functional, adequacy-based quality assessment models in use and being taught in translator training. If this study is anything to go by, I would consider usability methods a possible means to form new and applicable methods of TQ improvement and TQA in Translation Studies.
References

MATERIAL


LITERATURE


Rodríguez Rodríguez, B. 2007: Literary Translation Quality Assessment. Muenchen: LINCOLN.


INTERNET REFERENCES


**APPENDIX**

**Text excerpt**

**IMPROVISIINTI**


On kaksi olennaista tapaa paneutua improvisointiin sooloon: joko voit aloittaa soittamalla etukäteen määritellyä melodian ja jotka sitä etenemään koristellen sitä muistutteellaan sitten neuvostamo teeman säveltä ja soittotapana ei fra-seauraista tai sitten aivan saranmukaisesti aloittaa nollaasteesta vanoin tuen teella tai tietoisuudellaan siitä, mitä sitä soitaan tai jotenkö doottaa.

"Ainoa suunnittele meneen tehnnesi ja synty noita minuutti ennen soittamiseksi. Yrittä sitä välttää määrästään, mutta on koskaan josta suunnittelemaan siitä minuutti on."  

**Eric Clapton**

Improviseointi-termillä on laaja käyttö: säitä voidaan sovittaa joko soittomalla jokaisen soholin kokoan, jossa on sija, idearikas ja todella luova, tai sellaisena joka nojaa nippun vainokokoijiksi, jotka on harjoiteltu, palattu mieluummin ja lopuksi kytkeyt yhteeseen. Josain määrin useimmat kitaristit ovat kumpaakin laja.


**Jeff Beck**

"Minulla olik tapana harjoitella skaaloja, mutta ajattelun tavallisesti ottaa tarjonosan eri aseen avulla. Soitan juoksuutikasi jotka siirtyvät asemasta toiseen, joka harjoitellaan sooloon ja harjoitellaan melko helposti asemasta toiseen. Mutta harjoitellaan melko kehittämään ne suorat melko kehittämään ne suorat melko kehittämään ne suorat melko kehittämään ne suorat melko kehittämään ne suorat

**Albert Lee**


"Mahdollisimman monen kitarasoolon kuultavuus on parhaan menetelma alkuvaiheessa olevalle. Mutta saksafoni- sololikin vaatii opettajaharjoitteen. Ne ovat kilnastavia, koska niissä on myös kerrallaan vain yksi sävel, joka voidaan toistaa kitaralla. Jos osaat kopioida saksafonisoolon, voit jo hyvin, koska keskivertofo- nistis on jatkaa paljon paremmilta osaakin kitarista."

**Ritchie Blackmore**
Yksiääinen soolotekniikka


Näin yhden aseen muuntelutekniset on

Nauhaiskuteknikat (Slurrit)

Kaikki nämä keinot sisältävät kahden tai useamman sävelen soittamisen vain yhdel- la kielen näppäyskellä. Niihin kuuluvat ylöspäin nauhaiskus (hammer-on), nyhtäisy (pull-off) ja trillit.

Nauhaiskus


Yritä samaa 2. sormellasi 2. nauhalta (F, 3. sormella 3. nauheltal (G) ja 4. sormel- la 4. nauhastal Ost). Nauhaiskun vaihtus riippuu kahdesta

Nauhaiskut sommiteltu kielteita.

Nyhtäisyöt


Nauhaisku vapaalta kielelta.


Neljän sormen nauhanyhtäisyöt.

Trillit

Kahden sävelen nopeaa vuorottelua so- taan trillit! Se saadaan aikaan pihatankin sommitamalla alemmpi sävel, soitamalla se kerran ja sitten iskemällä sen yläpuolella nauhalla olevaa säveltä nopeasti ja tasa- sisti. Kielien pitäisi soida pitkään selvästi.
**Kielen venytys**

Venytys on eräs sõttoteknikan tärkeimpia aineikaa. Itse asiassa jokainen melodiasta tai valoosoittaa vastaa kitaristä käyttää sitä aina järjestyyn ja sitten on eräs nykykulttuurien tunnetuimpiin ominaispiirteihin.

Country- ja rockjamallinen kitaraponeerio James Burton pidetään yhtenä ensimmäisistä, joillakin kitaristina, jotka kiertävät kitarista ja sitten lopulta venytystä vastaavan musiikkiteknikkoja.

**Vibraatto**

Plenes "aaltolun" lisäämistä sävelseen saadaan vibraatoksi. Se merkitsee sävelkorkeuden nopeaa vaihtavaa muutetta. Vibraatto pidentää sävelen kestoa ja lisää ilmavisuutta sekä tallettaa yksinkertaisempikin meloodian tai soidun kuvon. Tätä tekniikkaa voidaan käyttää missä tilanteessa tahansa, mutta se on erityisen tehokas yhdistetyn kielten venytysten, venytystä kieltä sävelkorkeudelle ylös ja alas ja päädystä loppusuveluun sijoittaa siihen vibraatoksi.

Vibraatot aikaisemmin aukotuksissa havaitaan joko suoraan aukon tai kuuman aukon kanssa. Tällöin virheellisesti käytetään sanaa "vibraatoksi", kun se tarkoittaa ainoastaan kyseisen kielten ylä- tai alaspäin liikkumista.

**Liuku**

Liuku on vertailtavissa nauhaiskuun ja -nyh
tässyn sisällä, että vain yhdellä kielen näppäränliikkeen avulla saadaan soimalla usko perakkeliä saveltaja. Ero on siinä, että ensimmäinen ja viimeisen sävelen välissä jäävät sänet kuulillaan myös.

Liuku on helpompi soittaa kitaralla kuin ukeinmilla muilla soittimilla; joillakin se on täysin mahdotonta. Liuut on parasta soittaa kitaran nuoruus ja 7. sormella, koska tällöin 1. ja 4. sormi jäävät vapaillessa patka
taan soittaa jooppaampikumpaan suuntaan liuun lopputa.

Voimaa ovat nousevat liuut helpompia kuin laskevat. Noin niin, että koska liuku
kueen ylös kieltä pyrkii myös voimakkuus nousemaan. Tämäläisyyden lisäksi li turv
ta muita erityisasteja; esimerkiksi liu
tä säästää vajaaksi tai liuku
mattaa yhtä liikin

**Sormen puristus**

Sormen puristus on nousevien niin kuin laskevielkorikoin liukusvelen avais. Siinä on opittava loppusävelle liikkeen määrin sitä, että kättäytyneen sormin ainoa lujan kosketuksen piste. Kun sorm
ta soittaa, liuun päätösävyllä, pitääki
koko kavereen liikkeensä
riittävän

Tämä toimi liiurance

**Kahden nahtaven liuku**

Kahden nahtavan liukun on oltava suojelulla

**Täytän noinoinnutteen liuku**

Täytän noinoinnutteen liuku

**Vibsammon**

Vibsammon on vastaavan sävelkoneen nopeaa vaihtavaa muutetta. Vibraatto pidentää sävelen kestoa ja lisää ilmavisuutta sekä tallettaa yksinkertaisempikin meloon tai soidun kuvon. Tätä tekniikkaa voidaan käyttää missä tilanteessa tahansa, mutta se on erityisen tehokas yhdistetyn kielten venytysten, venytystä kieltä sävelkorkeudelle ylös ja alas ja päädystä loppusuveluun sijoittaa siihen vibraatoksi.

Vibraatoksi voidaan käyttää missä tahansa, mutta se on erityisen tehokas yhdistetyn kielten venytysten, venytystä kieltä sävelkorkeudelle ylös ja alas ja päädystä loppusuveluun sijoittaa siihen vibraatoksi.

**Vibsammon**

Vibsammon on vastaavan sävelkoneen nopeaa vaihtavaa muutetta. Vibraatto pidentää sävelen kestoa ja lisää ilmavisuutta sekä tallettaa yksinkertaisempikin meloon tai soidun kuvon. Tätä tekniikkaa voidaan käyttää missä tilanteessa tahansa, mutta se on erityisen tehokas yhdistetyn kielten venytysten, venytystä kieltä sävelkorkeudelle ylös ja alas.

**Kahden nahtavan liuku**

Kahden nahtavan liukun on oltava suojelulla

**Täytän noinoinnutteen liuku**

Täytän noinoinnutteen liuku

**Vibsammon**

Vibsammon on vastaavan sävelkoneen nopeaa vaihtavaa muutetta. Vibraatto pidentää sävelen kestoa ja lisää ilmavisuutta sekä tallettaa yksinkertaisempikin meloon tai soidun kuvon. Tätä tekniikkaa voidaan käyttää missä tilanteessa tahansa, mutta se on erityisen tehokas yhdistetyn kielten venytysten, venytystä kieltä sävelkorkeudelle ylös ja alas ja päädystä loppusuveluun sijoittaa siihen vibraatoksi.

Vibraatto voidaan käyttää vastaavan sävelkoneen nopeaan vaihtavaan muutettavaan. Vibraatomalla voidaan käyttää missä tahansa, mutta se on erityisen tehokas yhdistetyn kielten venytysten, venytystä kieltä sävelkorkeudelle ylös ja alas ja päädystä loppusuveluun sijoittaa siihen vibraatoksi.

**Kahden nahtavan liuku**

Kahden nahtavan liukun on oltava suojelulla

**Täytän noinoinnutteen liuku**

Täytän noinoinnutteen liuku

**Vibsammon**

Vibsammon on vastaavan sävelkoneen nopeaa vaihtavaa muutetta. Vibraatto pidentää sävelen kestoa ja lisää ilmavisuutta sekä tallettaa yksinkertaisempikin meloon tai soidun kuvon. Tätä tekniikkaa voidaan käyttää missä tilanteessa tahansa, mutta se on erityisen tehokas yhdistetyn kielten venytysten, venytystä kieltä sävelkorkeudelle ylös ja alas ja päädystä loppusuveluun sijoittaa siihen vibraatoksi.

Vibraatot aikaisemmin aukotuksissa havaitaan joko suoraan aukon tai kuuman aukon kanssa. Tällöin virheellisesti käytetään sanaa "vibraatoksi", kun se tarkoittaa ainoastaan kyseisen kielten ylä- tai alaspäin liikkumista.
noin, ettei se liu'u alta pois eikä tartu sormen lyhteen.

Uskunnaisia kitaroilta 2. ja 3. kielit ovat nampia joessa. Vakia, kielit voivat vaihdella molempien suuntien, antaa kielien työntämisen ylöspäin (baasoksi) tai kohti (yleen parreman tuntuman kuin vetäminen alaspaan (yläksi). Kun pyörittää kielitä ylös, on koko käsiesi sen takana. Kun vedät kieltä alas, on yhdem sormen tehtävä koko tyni.


Venytävällä sormella osuinen vahingossa muihin kieltei voi aiheuttaa tarpeet:

- Puolisivelaskeleen (1 nauha) venytys.
- Kokosivelaskeleen (2 nauhan) venytys.
- Puolentiosta sivelaskeleen (3 nauhan) venytys.
- Kahden sivelaskeleen (4 nauhan) venytys.

Tämä teknikka vaatii hyvaa kielien jännteet ja sävelkorkeuden tuntemusta, koska suunta on arvattava, joku pitkälle kieltä on venytettävä. Ei kuule sivettä ennen kuin se on jo suoritettu. Vastaavasti tähän ongelmään on harjoiteltu — ensin yhdem nauhan korkeutta vastaa venytys, sitten kahden ja kolmen — oppiakseesi tuntee enoten. Yritä sotita yksittäissävän ja venytä sitä sekä ylös että alas kahden tai kolmen puoliselaskelleen verran sen sollessa. Muista että kielten erilaisen jäkkyynyn vastaaminen "tässä" eli kokemus vaihtaa kieltä toiselle.

**Venytys sivullassa**

Kun olet oppinut haluttuseen venytyslevon, putoileeste sivelaskeleen sakska 3. sormelle, on aika oppia miten muita sormia käytetään. Kuten kalkessa kitaraekäskryssin, kyky käyttää yhte yllä hyvin kaikkia sor jokaisen parantaa sujuvuutta.

Keskisormesi ei pitäisi osoittautua liian harvoin, mikäli olet asettanut käsiesi niin että peukalosi edelleen antaa tukipiste tarvittavalle vipuvoimalle. Etusormella taittumen voi aiheuttaa ongelmia; ensiksi ki peukalosi tarjoama tukipiste ei ole enää tehokas; tekoiseen sinulla ei ole yhtä vapaata sormeaa kuolleammattia kieltä venytyspaikkaan takana; kolmannesta tarvitset tasaviallisesti etusormesi oikeonna muita sivelaja vähemmän joko ennen tai jälleen venytys. Kuitenkin aivan pitäisi kyettä venytämään etusormellesi ainakin puolisivelaskeleen verran 2. ja 3. kielillä. Pikkusormisen venytys ovat vaikeampia, koska pikkusormi on toisiaan harkimpa.

**Laskevat kielvenvenytys**

Suomenkielinen tiivistelmä

Johdanto

Nyt-liitteen artikkelissa Kustantaja, kirjassani on virhe! (HS, 44/2000) Tero Valkonen kritisoi huonoja englannista käännettyjä kirjoja suomenkielisessä käännöskirjallisuudessa. Yksi esimerkeistä on Ralph Denyerin Suuri kitarakirja (1982), jonka käännös on Valkosen mukaan ”varsinainen epäsuomen aarreaita”. Valkosen mukaan kirjan kielten huonot pitkään on niin huonoa, että ”lukija joutuu taistelemaan mahdottoman kieliasun kanssa päästäkseen perille itse asiasta”. Valkosen kritiikki toimii tämän tutkimuksen inspiraationa.

Tässä tutkimuksessa käytetään käytettävyystutkimuksen metodeja Suuren kitarakirjan arviointiin. Tutkimuksen kohteena on kirjan käytettävyys, mutta käsittelemme myös käännösten laadunaihe, joka on käännöstieteessä pitkään ollut ajankohtainen, mutta ongelmalinen aihe. Laadun määrittelyn ja arvioinnin hankaluuden tarkia tämän tutkimuksen fokus onkin käytettävyydessä. Käytettävyysmetodien soveltaminen on toistaiseksi melko uutta käännöstieteessä, vaikka käytettävyys itsessään ei mikään uusi asia olekaan.

Tutkimuksen ensisijainen tarkoitus on arvioida Suuren kitarakirjan käytettävyyttä käytettävyystutkimuksen metodeilla. Toissijainen tarkoitus on testata käytettävyysmenetelmien soveltuvuutta käännöksiin ja verrata niitä käännösten laadunarvioinnin teoreettisiin malliin. Tutkimuksen nimi perustuu Valkosen esittämään kommenttiin, jossa Frank Zappa kritisoi Elviksen kitaristien soittotyyliä ottamalla vertailukohdaksi Johnnie Watsonin ja Guitar Slimin kitarasoolot. Zappan kommentti kuuluu alun perin: ”That’s’ a guitar solo, nothing freeze-dried.” Valkosen mukaan suomenkielisessä versiossa kääntäjä ”on pannut parastaan” suomentamalla kommentin: ”Ne ovat kitarasooloja eivätkä mitään pystyynjääteníä kuivuuksia.”
Teoriatausta

Käytettävyys on lähtöisin ihmisen ja koneen vuorovaikutuksen tutkimuksesta (Human Computer Interaction studies), jossa se närhdään käyttöjärjestelmän ominaisuutena (Suojanen ym., 2012: 15). Käytettävyydessä päätavoite on käyttäjän kokemuksilla tuotteen käyttämisestä ja käyttäjä huomioidaan jo tuotteen suunnitteluvaiheessa. Standardissa ISO 9241-11 käytettävyys on määritelty seuraavasti:

Se, miten hyvin tietty käyttäjät voivat käyttää tuotetta tietyssä käyttötilanteessa ja saavuttaa määriteltyt tavoitteet tuloksellisesti, tehokkaasti ja miellyttävällä tavalla. [Suom. TS.]

Suojanen ym. 2012: 197


suoritetaan. Käytettävyyys käännöstieteessä voidaan nähdä pohjautuvan osittain funktionaalisiiin käännösteoriioihin, jotka keskittyvät käännöksen funktioon eli tarkoituukseen, toisin kuin perinteiset käännösteoriat, joissa tärkein ominaisuus on usein lähde- ja kohdetekstin välinen ekvivalenssi.

Käännösten laadun määrittely ja arviointi on ollut pitkään kiistanalainen aihe käännöstieteessä. Tämä tutkimus lähestyy aihetta käytettävyyden kautta, mutta on kuitenkin syytä tarkastella hieman miten laatua on määritelty käännöstieteessä ja etsiä yhtymäkohtia käytettävyyden arvioinnin kanssa.

**Materiaali ja metodit**


**Tulokset ja analyysi**

Heuristisen asiantuntija-arvion tulokset vihjaavat, että vaikka kirjan katkelmassa käytetty kieli ei ole kovin laadukasta, se ei kuitenkaan haittaa itse sisällön ymmärtämistä. Useimmat ongelmat liittyvät kömpelöön kieleen, kuten lauserakenteisiin ja sanavalintoihin. Esimerkiksi termin 'pull-off' kääntäminen termillä 'nauhanyhtäisy' vaikuttaa asiantuntijoiden mielestä erikoiselta tai jopa huvittavalta ratkaisulta. Asiantuntijoiden mukaan kirjan sommittelu on osittain ongelmallinen, sillä kuvat eivät ole kovin selkeitä. Informaation kulku nähäään kuitenkin loogisena ja luonnollisena, vaikka jotkin kohdat saattavat jäädä tulkinnanvaraisiksi. Huomionarvoista on, että arviossa ilmenneet kieleen liittyvät ongelmat olisi mitä luultavimmin voinut korjata käännettäessä, kun taas esimerkiksi sommitteluun voi olla vaikea puuttua käännettäessä.


Huomionarvoista onkin, että kun asiantuntija-arvioinnissa havaitut ongelmat painottuivat
kieleen, käytettäväystestauksessa havaitut ongelmat taas ovat luonteeltaan lähdetekstin ominaisuuksia, joihin voi olla hankala puuttua käännöspansessin aikana.

**Päätelmä**

Tutkimuksen ensisijainen tavoite oli arvioida *Suuren kitarakirjan* katkelmaa käytettäväysmenetelmillä. Tulosten perusteella käännöks ei välttämättä ole niin huono kuin Valkosen kritiikki antaa odottaa. Asiantuntija-arvion mukaan kieli ei ole kovin hyvää, mutta ei kuitenkaan niin huonoa, että se hankaloittaisi ymmärtämistä. Kääntettäväystestauksessa ei myöskään havaittu kieleen liittyviä ymmärtämistä haittaavia ongelmia. Suurin osa käytettäväystestauksessa löydetyistä ongelmista koski sommittelua. Tämä on mielenkiintoinen lisä perinteiseen käännösten laadunarvionäkökulmaan, sillä useat havaituista ongelmista olivat luonteeltaan sellaisia, joihin kääntäjä yksin ei välttämättä voi vaikuttaa käännöspansessin aikana.