UNIVERSITY OF VAASA

School of Marketing and Communication

Language Expertise in a Specialised Society

Heidi Ala-Kanto

Rookie Mistakes

Errors in Fan-created Finnish Subtitles in YouTube Videos

Master's Thesis in English Studies

Vaasa 2020

TABLE OF CONTENTS

PICTURES

TABLES

ABSTRACT

Tämän pro gradu -tutkielman aiheena ovat YouTuben suomenkieliset tekstitykset. YouTuben tekstitykset perustuvat pitkälti vapaaehtoisuuteen, videontekijöiden omaan aktiivisuuteen ja katsojien kiinnostukseen, ja niitä voidaankin pitää fanikääntämisen uusimpana muotona. Akateeminen tutkimus on tähän mennessä painottunut teksittämisen yhteisölliseen luonteeseen, mutta laadullista tutkimusta YouTube-käännöksistä ei ainakaan laajassa mittakaavassa ole tehty.

Tutkittavat tekstitykset kerättiin brittiläisen, Suomessa asuvan youtubettaja Dave Cadin videoista, koska hänen materiaalinsa oli minulle jo ennestään tuttu ja hänen kohdeyleisönsä on lähes kokonaan suomenkielinen. Tekstitykset koottiin 17 videosta, ja niille suoritettiin virheanalyysi. Koska tekstitysten laadun tutkimiseen ei ole virallisia ohjeita, muokkasin tutkimukseen sopivat kriteerit suomen kielen kieliopin sekä suomalaisten tekstitysstandardien pohjalta. Virheet luokiteltiin kielioppi-, käännös-, oikeinkirjoitus- sekä teknisiin virheisiin. Tekstityksissä oli virheitä kaikista näistä kategorioista, mutta varsinaisia kielioppi- ja käännösvirheitä oli odotettua vähemmän. Sitä vastoin oikeinkirjoitusvirheet, etenkin puuttuvat pisteet virkkeiden lopuissa olivat tutkimuksessa yliedustettuina. Teknisistä virheistä korostuivat liian pitkät tekstitykset, mutta ne olivat amatöörikääntäjiltä odotettavissa.

Tutkimukseni osoitti, että toisin kuin ehkä voitaisiin olettaa, amatöörikäännösten ongelmat eivät korostu käännös- tai kielioppivirheissä vaan oikeinkirjoitusvirheissä. Osasyynä tähän saattaa olla Internet-viestinnässä, joka etenkin englanninkielisessä viestinnässä suosii välimerkitöntä kirjoitusmuotoa, jossa usein myös isot alkukirjaimet jäävät pois. Tekniset virheet taas selittyvät amatöörikääntäjien tietämättömyydellä. Esimerkiksi liian pitkiä tekstityksiä voitaisiin vähentää lisäämällä YouTuben tekstitystyökaluun ohjeistuksia merkkimääristä.

1 INTRODUCTION

Every generation has its innovations and phenomena that become their defining characteristics and a part of their identity: Baby Boomers, born after the Second World War, have the hippie movement and prosperity. Generation X, born between the early 1960s and early 1980s, has liberal parenting and the MTV. Generation Y, born between the early 1980s and early 2000s, and Generation Z, born after mid-1990s, both have the Internet. Most of Generation Y, perhaps better known as Millennials, and practically all of Generation Z have a hard time recalling a time when the Internet was not a part of our everyday lives. The modern society has grown dependent of it, so much so that it can cause problems to not have access to it, for example when bank services are moved exclusively online.

One such "problem" that the popularity of the Internet as a platform has caused is the decreasing popularity of television as a media. According to a 2017 study, 47% of Millennials and Generation Z are "unreachables", people whose media consumption cannot be tracked because their watching habits fall outside the traditional tracking methods (Heart & Science 2017). However, not only does this reveal a change in watching habits, it also shows that the research methods are irreversibly outdated and quickly becoming obsolete.

The same study (Heart & Science 2017) revealed that 73% of Millennials use online streaming services such as Netflix, Amazon, Hulu and YouTube. The first three are used to stream series and films created by massive film studios and production houses, whereas YouTube is a more diverse platform with its content created by ordinary people. Over a billion users use YouTube to watch and upload videos, everything from funny cat videos to educational content and personal video blogs, also known as vlogs (YouTube 2018). The content creators can be individuals, corporations, or anything from between, which makes YouTube stand out from the traditional media: for example, vlogs bring the content creator close to his or her audience and create a new kind of a "one man show", thanks to the interactive nature of the platform.

The internet has made the world seem smaller, and while every country has their own YouTube celebrities and viral videos, especially the English-speaking YouTubers – people who make videos on YouTube – have the possibility to amass an international following of millions. While it is true that English has become the lingua franca online, not all Internet users, also known as Netizens or Net Citizens (Hauben 1997: 3) speak or understand it fluently enough to watch English-speaking YouTubers without effort, not to mention the special needs of the deaf or hard-of-hearing, which is why YouTube has integrated a subtitling tool in the videos. The interest of this study lies in the fact that these subtitles are almost exclusively made by members of the audience and are essentially fan translations. Academic research on YouTube translations so far is mainly focused on the aspect of co-creative labour and participatory culture (e.g. Banks & Deuze 2009 and Dynel 2014), instead of studying the people participating in amateur translation or analysing the translations as a type of amateur translations.

The aim of this study is to research the overall quality of Finnish YouTube subtitles made by fan translators. The main focus of the analysis will be on translation errors: what kinds of translation errors can be found in the study material, whether there are recurring mistakes, as well as differences from the Finnish subtitling customs, even though it is not assumed that amateur translators are actively aware of them. I am also interested in whether the fact that they are being made by amateurs is visible in them, for example through a common pattern or similar translation mistakes or solutions. My research questions for this study are: 1) What types of translation errors can be found from the Finnish YouTube subtitles? and 2) Are there translation errors that recur in these subtitles?

The videos on which I am basing this study are made by a British-born YouTuber Dave Cadwallader who goes by the username Dave Cad. What makes his videos suitable for this study is the fact that, even though he speaks English, his target audience is Finnish-speaking, which means that the subtitles are made by Finnish audience members. It is therefore easy for me to analyse the Finnish subtitles, spot the mistakes and deduce the thinking behind different translation solutions.

I will begin my thesis by introducing AV translation, its history and developments briefly in Chapter 2, focusing on subtitling and its limitations. I will also discuss translation error analysis and how it will be adapted and employed later in the study. Then, in Chapter 3, I will discuss YouTube, its translation tool and its connection to machine translations, as well as fan translations. Chapter 4 will be dedicated to the analysis of the material in which I will introduce, analyse and discuss the translation errors found in the Finnish YouTube subtitles. The conclusions in Chapter 5 will include a description of the research process, discussion on the findings of this study and its limitations and, finally, suggestions for further studies.

1.1 Material

The primary material for my thesis will consist of two groups: firstly, I will conduct a questionnaire-based survey among YouTube translators. I contacted the translators who had allowed their name to be visible as the translator of a video and sent them a direct link to the questionnaire form if they had some contact information listed in their profile. In total, the number of translators I contacted was 5. The questionnaire was conducted on Google Forms and the questions focused on the backgrounds of the translators: how old they are, what their occupation and education are, have they studied translation or languages in general, do they have previous experience in translation or possibly a multilingual background. The aim is to find if there are common nominators in the fan translators' backgrounds, to see if perhaps these nominators are what have encouraged these people to translate the videos. The translators will submit their answers anonymously.

Secondly, for the main part of my primary material, I will select some of Dave Cad's subtitled videos and focus on their subtitles. In November 2017, Dave Cad posted videos three times a week, which means that the amount of potential subtitled videos keeps growing every week. On November 2nd, 2017, when this study begun, Dave Cad's channel had 136 videos, of which a little less than half, 53, had Finnish subtitles. Out of

those 53 videos, 37 had translators who had agreed to share their name in the video's description.

Considering the length of the thesis, I had to limit the number of videos included in the analysis. Hence, I would only include videos with a visible translator and only include one video per translator, even though the subtitles would not be analysed in connection to the translators. This criterion was met by selecting the first video with a new translator's credits, counting from the oldest videos towards the newest so that I would affect the selection process as little as possible. By limiting the number of videos based on the translators' visibility, I hoped to minimise the amount of translations made by multiple authors. It is still possible that other people have contributed to the subtitles, but it could be argued that people are more likely to create subtitles to videos that do not have existing subtitles or translators. There were also a few videos where there was visibly more than one translator: these videos were excluded from the analysis.

Based on these criteria – a visible translator, one translator per video as well as one translated video per translator – the total number of subtitled videos included in the analysis is 17. The subtitles from these videos were copied into a Word-file on April 11th, 2018, to avoid a situation where subtitles are edited while the research is still ongoing. In case editing has occurred, the analysis will be based on the version in the Word-file. In total, the subtitles amount to about 15,200 words, and the total length of the videos amounts to 123 minutes and 11 seconds.

1.2 Method

The research method of this thesis will be an evaluative case study employing mixed methods. The focus of this study is on the amateur-made subtitles and error analysis. Since the videos are being translated by different people, the focus of the product analysis will be on translation errors in general. Different types of translation errors will be collected, categorised and analysed based on a customised model, as there are no definite guidelines or instructions for creating or analysing subtitles. However, there have been

attempts at creating quality assessment models have been made, e.g. Pedersen's (2017) FAR model and the recent Finnish subtitling quality recommendations (2020) created by a group of translators and representatives from translation houses, TV channels, streaming services and Kotus (The Institute for the Languages of Finland).

The FAR model in turn is inspired the NER model, a quality assessment model for intralinguistic translation for the deaf or hard-of-hearing. Petersen's model uses three distinctive areas in its assessment: functional equivalence, acceptability and readability. (Petersen 2017: 217) Before discovering Pedersen's FAR model, I created my own assessment model which shares similar elements with it, but instead of the abovementioned 'umbrella' categories, the model used in this case study is built using four more definitive and specific categories, introduced in Chapter 4. The model I created is based on the Finnish subtitling conventions, as well as the grammar rules of the Finnish language. Subtitling and its conventions will be discussed in section 2.1, and the complete model will be introduced in the analysis section 4.2 of this thesis.

This study will mainly consist of qualitative product-oriented research, namely analysis on the quality of YouTube subtitles. Quality in this thesis is more focused on the different types of errors that can be found in them. A quantitative element is added by counting and categorising those errors, as well as seeing how the errors are distributed across the different categories. Texts, in this case audiovisual texts, are a popular research subject for translation studies and thus also for case studies written on translation. The research process will follow an iterative pattern, as some analysis will happen simultaneously with the data collection. (Saldanha & O'Brien 2014: 118, 122) This will most likely happen when the research material, i.e. the videos and their subtitles, is transcribed and copied, both for preservative purposes and easier analysis. This also applies to when translation error analysis is conducted on the subtitles.

In addition, a questionnaire-based survey will be conducted to receive information on the translators. Matthews and Ross (2010, quoted in Saldanha & O'Brien 2014: 85) define a questionnaire as "a list of questions each with a range of answers" and "a format that enables standardized, relatively structured, data to be gathered about each of a (usually)

large number of cases". It can be used to gather information on the research participants' background or opinions, behaviour, etc., and it is perceived as the easiest tool to analysing large quantities of data. As with any research tools, a questionnaire also has weaknesses: for example, the design can be flawed, there are too few participants to receive a trustworthy sample of answers, the research participants may 'sabotage' the answers by answering untruthfully or their answers might be affected by the research situation, i.e. knowing that they are participating in research might make them choose "nicer" or more flattering answers. (Saldanha & O'Brien 2014: 86)

Internet-mediated collection method was chosen as the data collection method as it would have been impossible to contact the research participants in any other way. Conducting the survey online also makes it easier for the participant to answer the questionnaire regardless of their location. (Saldanha & O'Brien 2014: 92) Google Forms was selected as the questionnaire platform as it is both recognisable and simple, and it is possible to submit an anonymous answer. It also has a summary tool that shows the answers to the close-ended questions in a ready-made pie chart. A pilot test for the questionnaire used in this study was conducted to test the platform as well as the questions and based on the pilot testers' feedback some of the questions were modified to avoid ambiguity.

The questionnaire consists of three separate sections: the first section has questions related to the translators' background and the second section has questions on their language proficiencies. The third section of the questionnaire has a series of questions related to the translation process of YouTube videos and possible difficulties the translators might have faced. Section 4.1 of this thesis will elaborate on the questions in more detail.

The research participants, the amount of which was finally reduced to 5 translators from the original 18, were chosen based on non-probability sampling which means that they were the easiest to contact (Saldanha & O'Brien 2014: 91). In order to find the translators who could potentially participate in the survey, I first had to determine which of the 136 uploaded videos uploaded on Dave Cad's YouTube channel before November 2017 had Finnish subtitles. Next, I had to find the translators' usernames and links to their accounts

in those translated videos, and out of those users I had to eliminate the ones who had no contact information available.

The link to the questionnaire was sent from my personal accounts to the 5 translators on October 1st, 2018 via Gmail or Facebook, depending on which contact information was available. The response rate was expected to be quite low, as is often the case with internet-mediated questionnaires (Saldanha & O'Brien 2014: 92). In the end, three participants responded: it is not reliable to make generalisations based on only three answers, but the answers will be presented and discussed as an added interest. The questionnaire, as well as its English translation, is included in the appendices. The results for the questionnaire will be discussed in section 4.1, and they will be presented in full in the appendices.

The theoretical framework for my thesis will be heavily based on the theories present in the field of audiovisual translation, referred also to as AVT or AV translation. In addition, translation error analysis will be used to support my own analysis on the translation errors found in YouTube subtitles. I will also discuss subtitles in general and fan or amateur translations, as that is what the subtitles on YouTube videos are. However, there does not seem to be any studies on YouTube subtitles or YouTube translators, specifically. For both AV and fan translations, I will refer to such scholars as Díaz Cintas, Pérez-González, and Susam-Saraeva.

1.3 Dave Cad

Dave Cad is the pseudonym of the British-born YouTuber Dave Cadwallader. He has made videos on YouTube since May 2011 and by April 15th, 2020, his channel had approximately 145,000 subscribers, with in total almost 25 million views on his videos. What makes Dave Cad's videos an interesting and optimal subject for this translation research is his demographic: most of his audience is from Finland and Finnish-speaking, even though he himself does not speak Finnish. His popularity among Finnish viewers is certainly due to the fact that most of his content is about Finland, Finnish culture and

Finnish food, especially now that he lives in Helsinki, Finland with his family. It could also be argued that the demographic and content of his videos contribute to each other.

Most of the subtitles that the videos have are in Finnish, some are in English, and there are a few videos where the subtitles are in some other language, such as Danish or Icelandic. Since the amount of subtitling languages is limited, it also makes the research of the translators easier and more thorough, compared to other English-speaking content creators with viewers from all around the world and subtitles in dozens of languages.

YouTube is a video-publishing platform that allows its users to upload and watch online videos free of charge. It is an enormous community that connects content creators and viewers on both local and international levels. Content creators are free to upload whichever type of content they prefer, as long as it does not violate any of YouTube's Guidelines. Dave Cad's content is varied, but in its core, it is based on vlogging, short for video blogging. In these video blogs, referred to also as vlogs, the content creator is usually in the centre of the video, just talking to the camera about a personal experience, opinions on current themes or updates on their lives. Daily vlogs are a type of vlog where the camera follows the content creator throughout the day, and it may include fillers, such as a montage of short video clips of the environment, other people, etc.

As with for example fashion, YouTube also has trends. Such trends have been for example test videos where people try food or other products on camera and give their opinion on them and react videos where they film their own reaction to other videos or phenomena. Q&A videos, that is, videos where viewers have submitted questions and the content creator answers them are also popular. Dave Cad has for example made videos of him trying Finnish candies and reacting to Finnish music videos.

English-speaking channels naturally have the potential to reach huge audiences, whereas channels in other languages, for example Finnish, have a very limited reach. Adding subtitles to their videos adds to the content creators' potential to amass international audiences, and it also gives people with hearing impairments a chance to access their content, presuming that there are subtitles available in their native language or any

language they understand. The content creators may add the subtitles themselves, depending on their language skills and preferred target audiences, or they can ask their audiences to contribute.

For videos in English, it is also possible to add auto-translations generated by a machine which automatically detects the language used in the video and adds subtitles. The tool is still far from perfect, but it can be utilised in making subtitles in English for videos in English. YouTube and its translation tool will be introduced in more detail in Chapter 3 of this thesis. The following chapter will introduce audiovisual translation, with the focus being on subtitling. Machine translation will be briefly discussed, after which translation error analysis will be introduced and discussed in connection with subtitling.

2 AUDIOVISUAL TRANSLATION AND TRANSLATION ERRORS

Audiovisual translation, often shortened as AVT, has emerged and will continue to develop following technological achievements and discoveries. Despite audiovisual translation existing since the emergence of film, it has taken some time for the academic circles to notice it: AVT research has become a significant area of study only after the turn of the millennium (Pérez-González 2014: 12). Nowadays, audiovisual content and therefore also audiovisual translation are an inseparable part of our everyday lives: whenever we turn on the TV, computer, or our smart phone, we are met with an overload of pictures and sound in different forms of advertisements, programmes, or applications. The academic interest in AVT research is also visible in the number of conferences held and theses and dissertations written on the subject, as well as journals and periodicals (Gambier 2008: 14).

Audiovisual content, or audiovisual texts are defined by four basic elements: the acoustic-verbal, the acoustic-nonverbal, the visual-nonverbal and the visual-verbal. The acoustic-verbal elements include everything that is spoken language, for example dialogue and songs, whereas the acoustic-nonverbal elements include sounds effects and music without lyrics. The visual-nonverbal elements are also without language, for example pictures, paintings or gestures, and visual-verbal elements in turn include language, for example in signs, inserts or letters. Out of these categories, it would seem that the visual-nonverbal is the most important one. (Delabastita 1989, quoted in Díaz Cintas 2008: 3) It could be explained by the fact that images are to a large extent universal, and non-verbal messages are more effortlessly transmitted and understood by a large and diverse demographic than, for example, complicated dialogue.

Audiovisual content was first realised in the form of cinema in 1895, when silent films were invented. Dialogue was presented in text form between scenes, known as intertitles, and music was often provided by live performers. Sound became an irreplaceable part of film in the late 1920s when talking films were introduced to the audiences. This development, and the film makers and producers' want and need to reach new audiences all over the world created the need for audiovisual translation. Subtitling and dubbing

were one of the first forms of audiovisual translation that are still used: in addition, multi-lingual versions of films were produced, meaning that the film would be shot multiple times with the actors performing in different languages. (Díaz Cintas 2008: 1–2) Multi-lingual translation method soon vanished as it was both expensive and time-consuming to re-shoot the films.

Not only have new forms of AVT been developed in addition to subtitling and dubbing, but also new media incorporating AVT have emerged. Television was the next important medium in the development and distribution of audiovisual content. Television has always been and will continue to be an international medium (Immonen 2008: 8) that brings the world's events and different cultures to the viewers' living rooms. The broadcasting of the first moon landing in 1969 was an event that glued people around the world in front of their TV sets and undoubtedly proved the power of television as a medium. Even to this day, television is an efficient channel to reach massive audiences, even though the Internet will most likely de-throne it as the go-to medium in the coming decades. The medium and its content might change, but the need for translation will not disappear.

AVT research, as stated before, has become increasingly popular in the 21st century, and there are several catalysts that lead to the rise of AVT research. One of the first ones was the year 1995, and the celebrations for the hundred years of cinema. Before the mid-90s, AVT had not been systematically studied, and the research that had been conducted did not spark interest among researchers. Other factors that made the 1990s significant for AVT research were migration and technological developments. (Gambier 2008: 12) The changing geopolitical situation in Europe, as well as developments in Asia and Africa resulted in increases in immigration and the need for translation and language teaching.

One of the key developments in technology was digitalisation, which began in earnest in the 1990s. Audiovisual content was previously produced for devices using analogue technology, whereas now new devices employing digital technology started conquering the technology. Digital technology was not only faster, but information also took less space, which made, for example, DVDs an extremely popular format for storing

audiovisual content. (Gambier 2008: 25–26) DVD added new possibilities to subtitling and enabled multiple subtitling tracks to be added on a single disc, which was very cost-efficient, but it also resulted in an increase in intralingual translation. This was significant especially in countries where dubbing was the preferred method and as a result the deaf or hard-of-hearing had previously had little to no access to audiovisual content. (Díaz Cintas & Remael 2007: 17–18)

With the development of the field, AVT has become more than just translating audiovisual texts into other languages in order to make foreign films and TV programmes understandable to audiences across language borders. For example, AVT also has a new role in language learning for immigrants. (Díaz Cintas 2008: 6) Watching TV with the audio and subtitles both in the source language, for instance Finnish, helps people with different language backgrounds both hear and visualise the new language.

As AVT has gained more popularity, also the tools for audiovisual translation, such as subtitling programmes have become easier to use and accessible for everyone: some of the programmes available are even free to use, which in turn has encouraged the tradition of fan-made subtitles (Díaz Cintas 2008: 7). The main interest of this study is in fan-generated subtitles on YouTube videos, which is why I will be focusing on subtitling, its conventions as well as limitations in section 2.2. It needs to be noted that subtitles and the conventions relating to subtitling will be discussed from the Finnish point of view. In section 2.3, I will introduce and discuss translation error analysis and how it can be connected to subtitling. Fan translation and its conventions will be discussed further in Chapter 3.

2.1 Subtitling

There are multiple ways of translating audiovisual material, but the three main ones are dubbing, voice-over and subtitling (Díaz Cintas & Remael 2007: 8). Subtitling is the act of presenting a translation written in a target language in synchrony with the corresponding line of dialogue spoken in a source language (Pérez-González 2014: 15–

16). Dubbing and voice-over preserve the form of expression, that is, sound, whereas subtitling adds the element of text. The interest of this study lies specifically in subtitling.

A subtitled programme consists of three elements: speech, the image, and the subtitles, with the subtitles usually situated in the lower part of the screen. Since the medium, for example film, sets spatial and temporal limits for translation, some do not acknowledge subtitling as translation proper: instead, they talk about it as adaptation or rewriting. Indeed, the subtitles may differ from the content of the source language and for example omit or shorten it to respect the principle of synchrony, so much so that it can no longer be regarded as a translation. (Díaz Cintas & Remael 2007: 8–9, 144–145; Immonen 2008: 10) However, subtitling also incorporates the source text in the final product in a way that 'traditional' translation does not, which allows the translator to make perhaps more liberal translation decisions.

Every type of translation has its constraints and problems, whether they are related to the product, medium, or external factors such as production timetables. Every final text is produced after reading, interpreting, and choosing, and whether that text is good or bad cannot be judged out of context as it is always affected by multiple factors. (Díaz Cintas & Remael 2007: 145) Subtitled texts should therefore be assessed in its own context and avoid comparing them to, for example, novel translations in terms of omission, expression, etc.

Time is an important factor in subtitling as well as in any form of audiovisual translation also from another perspective: after a film or a TV show is finished, subtitles need to be added before the final product can be distributed, and the time allocated for the actual translation work is often extremely limited. Regarding this time-aspect, subtitles can be divided into pre-prepared and live or real-time subtitles. Pre-prepared subtitles, the process also known as offline subtitling, are done after the audiovisual product is finished, whereas live subtitles, in a process called online subtitling, are added as the product is being broadcast. (Díaz Cintas & Remael 2007: 19)

In offline subtitling, the translator has more time to refine and adapt the subtitles to the product. These subtitles can be categorised based on their lexical density. In other words, pre-prepared subtitles can be further divided into subtitles with reduced or complete sentences. Live subtitles require quick reaction, and there is no time for editing. Live subtitles can be categorised as human- or machine-made translations. (Díaz Cintas & Remael 2007: 19) Pre-prepared subtitles are the traditional form of subtitling, but as the technology evolves, live subtitles with the help of machine translators can be expected to become more common as well as more accurate.

Subtitling is the most common mode of audiovisual translation in Finland: roughly 80% of foreign-language programmes on YLE, the national-owned broadcasting company, are subtitled into Finnish and/or Swedish. The technical limits set by the medium force the translators to sometimes condense the subtitles considerably compared to the source text. According to the subtitling conventions on YLE, the subtitles must fit on two lines with the maximum of 35 characters per line on the bottom of the screen, and they appear for 1,5–6 seconds depending on the length of speech or scene. The change of speaker is indicated with a dash and a space (-) in front of the line of dialogue. (Immonen 2005: 167, 171) The technical aspect of the subtitle analysis in Chapter 4 will be based on these criteria, mainly on the length and duration of the subtitles. It is unlikely that amateur translators would make the subtitles based on them, as they might not necessarily be aware of the rules, which is why it is assumed that there will be several instances of technical translation errors.

Creating good-quality subtitles in the given time and space can be quite difficult, especially since the words and sentence structures in Finnish tend to be significantly longer and more complex than, for example, in English. Because of this, the translator needs to extract the main points of the dialogue and inevitably omit some content. As the original audio is still audible to the viewers, it can be assumed that the audience gets clues from it, for example through pauses and facial expressions. Cultural references are often difficult to translate or explain, but the more distinctly different culture is presented in the AV product, the more easily the differences are accepted by the audience, as opposed to,

for example, through dubbing, where the clues from the original audio are erased. (Herlin 2008: 137–138)

In other words, cultural differences are more easily explained through subtitling, as the viewers can also utilise their own cultural knowledge. Based on this, it could be argued that subtitling gives the audience the possibility to analyse the original product and its content themselves, whereas in dubbing the audiovisual product the responsibility on the understandability relies solely on the production and dubbing team.

The research of interlinguistic subtitling has traditionally received most of the researchers' attention in the field of audiovisual translation, and more specifically, the interest has been in the study of differences between the source and target texts. The study of interlinguistic subtitling can be divided into three distinct categories: case studies focus on certain audiovisual texts and products, whereas some studies focus on specific issues in audiovisual translation, such as the translation of humour or cultural references. There are also studies that focus on subtitling strategies. (Gambier 2008: 17–18) In light of these categories, this study as a case study belongs in the first category, even though it also combines the study of the producers or participants, in this case fan translators, which are not mentioned in Gambier's categories.

Audiovisual translation is a field with multiple participants and intense competition: there are commissioners, producers of the audiovisual content, translators, and editors to just name a few. Commissioners want to get subtitles made as cheaply as possible within the production timetable, while translators fight to receive an adequate compensation for their work. Both fan translators and machine-generated translations add to this division, and fan translators' role and effect on the field will be further discussed in section 3.3. Machine translations are possibly the latest development in the field of audiovisual translation, or at least it is starting to gain recognition as a valid tool for translation.

As technology advances, tools like machine translation programmes become better utilised and their quality increases. For now, it would seem that machine translators work best as a tool for the human translator, not as an independent translator itself (Linna 2013:

7), but with researchers, scientists, engineers and programmers constantly working on devices, programmes and applications, it is not an impossible thought if in a couple of decades machines master human languages. Already, there are smart phone applications made for tourists that can detect and translate text from pictures and automatically detect and translate speech (Telegraph 2018).

Machine-generated subtitles are also a reality on YouTube, but for now, they are available in only a couple of languages: English, Dutch, French, German, Italian, Japanese, Korean, Portuguese, Russian and Spanish. YouTube uses speech recognition technology to detect the language and automatically generate subtitles, but the technology is still prone to mistakes. (YouTube Help 2018) However, compared to for example Finnish, English is a relatively simple language for machines to interpret: sentence structures are straightforward, there are significantly less inflections, and nouns do not have genders like, for example, in Russian or German.

The abovementioned subtitles are intralingual, which means that they translate spoken language into the same language in written form. Even though machine translations currently work best with the English language, there is still a long way to perfect machine-generated translations in any language. However, even the imperfect translations can be utilised by human translators: they may require varying amounts of editing, but they can lighten the translators' workload, shortening the translator's work process possibly by several hours. The same can be applied to YouTube translations. YouTube's translation tool will be introduced in detail and with examples in section 3.2.

As stated, there is still plenty of room for improvement in machine-generated translations. Machines, as well as humans, make mistakes on different levels of translation, from typing errors to context-related mistakes. The following section will discuss translation error analysis and the different qualifications for a good translation, which will form a base for the analysis section of this thesis.

2.2 Translation quality and error analysis

Defining a good translation is difficult, but one of the main requirements for a good translation is that the language is grammatically correct and that it follows the linguistic conventions of the target language, whether it relates to conjugating words of foreign spelling or something as simple as punctuation. As unfortunate as it is, mistakes evoke a greater reaction than success, and such is the case also in translation: if the subtitles of a film are well done, they are effortless to read and seem like a natural part of the final product, but even the smallest of errors draw the viewer's attention in a negative way. Acknowledging different translation issues could help improve the overall quality of the translation and minimise the amount of translation errors. The following translation issues are not solely applicable to subtitles nor are they limited to these examples, but they are significant to subtitling.

Firstly, the translation of linguistic variation and marked speech: there are differences between spoken and written language, the way people use language depending on their socio-cultural background, there are different dialects, sociolects and idiolects, and even accents that all affect the way people speak (Díaz Cintas & Remael 2007: 184–185, 187–195). On these occasions, the translator is faced with multiple choices: will they try to imitate the style, dialect or accent present in the original text, will they try to find a similar option from the target language, or will they ignore it altogether, instead focusing on the content?

Secondly, translators may face issues with single words, namely those with connotative meanings, and powerful words such as taboo words, swearing and interjections, as well as culture-specific words (Díaz Cintas & Remael 2007: 185–186, 195–207). The translator needs to be aware of the connotative meanings in the original text and decide whether to try replicating the same in the target language or not. The same applies to taboo and swear words: with these words, the translator also needs to remember the target audience, especially with swears. Translating culturally specific terms can be resolved by using, for example, loan words or similar terms. Explaining the term is effective but difficult, given the limited time and space of subtitles.

Thirdly, the translator may have to work with a text that includes songs. The first decision to make is whether the song needs to be translated in the first place: is it vital for the audience to understand the lyrics or is it just an artistic addition? The song might also be famous enough to be left untranslated based on the assumption that the viewers will understand what it is about. If the song needs to be translated, priority needs to be given to either content, rhyme or rhythm. (Díaz Cintas & Remael 2007: 207–211) Sometimes, especially with older songs, the lyrics already have official translations. If that is the case, the translator could rely on those if they are suitable for the medium and they have a permission from the right's owner if it is not in the public domain.

Fourthly, another source for translation issues is humour. What is funny in one language and culture may not be amusing in another. The translator needs to recognise what the source of the humour in the text is: is it wordplay, political or social commentary, an obscure reference, or some other context? Can the joke be transferred so that it originates from the same source, or does it have to be localised or changed altogether? (Díaz Cintas & Remael 2007: 212–228) Translating humour is difficult, especially since everyone has a different insight as to what they think is amusing.

Finally, sometimes resolving translation issues can lead to choices that are not as neutral as they should or could be. It is impossible to achieve a completely neutral translation, as choices and decisions must be made during the translation process. These can cause ideological issues, especially with texts that consist of multiple languages in an area where one language is dominant. For example, in a situation where speech is overlapping, the dominant language may be given more authority over a minority language. (Díaz Cintas & Remael 2007: 229–232) All of these issues demonstrate just how difficult translating can be and how aware of their own decisions and their effects translators must be, and sometimes they do fail in acknowledging different problems. Ignoring these abovementioned issues or failing to resolve them can lead to a sub-par translation or even a translation error.

There are different ways of defining a translation error, but generally it refers to an anomaly, a deviation from a translation convention. An error can be a blatant mistake,

such as a spelling mistake, but sometimes also a poor translation can be counted as an error, depending on the evaluation criteria. Defining a translation error also depends on the translation mode in use and, for example, what type of equivalence needs to be employed in the translation, if any. The aim of formal equivalence is to produce a similar form to the source text, whereas semantic equivalence focuses in replicating the content of the text. A dynamic equivalence aims to reproduce the effect of the source text which sometimes requires changes in both form and content. (Vehmas-Lehto 2005: 49–50) Formal equivalence is possibly most important when translating poetry, and dynamic equivalence is a priority when the aim is to evoke an emotional response. Semantic equivalence offers the translator more flexibility and artistic freedom regarding the form of the text but at the same time the responsibility to produce equivalent content increases.

In a short case study, Łukasz Boguki analysed subtitling errors made by a Polish amateur translator. The translator did not have the original source text in a written form when creating the subtitles: instead, they worked with a poor-quality recording of a screening in cinema, which also affected the final product. Based on the errors found in the subtitles, Boguki lists five factors that caused erroneous subtitles. Firstly, some errors were made due to the translator's insufficient language skills in that they did not recognise some of the less common vocabulary, combined with the inability to see the words in written form. Secondly, there were sentences that were not completely understood, and the translation decisions made based on misunderstood information resulted in errors. Thirdly, errors were made based on misinterpreted ellipses: for example, "few" was interpreted as "a few". Fourthly, there were words and full sentences that were misinterpreted to mean something else based on hearing alone. Finally, there were errors that were a result of an overreliance on context, for example visual cues. (Boguki 2009: 49–57)

Vehmas-Lehto defines translation errors based on six categories: the severity of the error, the "level" where the error can be found, the translation phase where the error has occurred, the cause of the error, the source and target languages, and overt and covert errors. The first category, the severity of the error defines the effect the error might have on the text, its content and readability. Level in the second category refers to errors relating to either linguistic or textual and extratextual factors. In the third category, the

error can occur in three different phases: analysis, transfer and format phase. An error in the analysis phase is most likely due to a misunderstanding or misinterpretation, in the transfer phase it is often caused by interference when the text is converted from source language to target language. Errors in the format phase are deviations from the external requirements set for a target language text. (Vehmas-Lehto 2005: 53–69)

There are several reasons in the fourth category as to why errors may occur. Poor language skills can cause mistakes in the understanding or formatting of texts, but there are also errors that derive from the translation process and its "special nature". The lack of cultural knowledge and the weak "contrastive competence" of the translation are causes for such errors. Recognising cultural significances and nuances in meanings in the source language helps the translator understand what he or she needs to do in order to make it accessible in the target language. It also helps the translator to avoid interference when he or she has the tools to recognise even the slightest differences between source and target languages. It is worth noting that only the unconscious influence from the source language can be categorised as interference. (Vehmas-Lehto 2005: 56–57)

The fifth category for translation errors is based on the translation languages, that is, the source and target languages and the translator's language skills. Linguistic errors are more common when translating into a foreign language, whereas translations from a foreign language have more traits of interference from said foreign language. (Vehmas-Lehto 2005: 63) It needs to be noted that in this case, it is assumed that either the source or target language is the translator's first language.

The sixth and final category of translation errors divides them into two groups: overt and covert errors depending on their noticeability. Overt errors are easy to locate and clearly identifiable as mistakes: they include deviations from the content of the source text and deviations from the linguistic conventions of the target language. In other words, they alter the content of the original text or interfere with the readability of the target text. Covert errors are less obvious but still identifiable as mistakes: a text with covert errors might at first seem to follow linguistic conventions but proves difficult to read due to unorthodox translation solutions or interference. (Vehmas-Lehto 2005: 64–65)

The error analysis conducted in this study will focus on both the content and the form of the subtitles. Vehmas-Lehto's second and sixth category will heavily inspire the so-called 'error categories' constructed and employed in Chapter 4. It is assumed that YouTube translators are not consciously following some translation conventions or strategies and are instead unconsciously aiming to produce a semantically equivalent translation, as that suits the medium, but it is also the form of translation that laymen might connect to translation altogether. Interference is also expected to be found in YouTube subtitles, but its cause can be difficult to locate, as it can originate from multiple sources as presented in Vehmas-Lehto's categories.

Subtitles in YouTube videos can be classified as fan translations, and both YouTube and amateur translations will be introduced next in Chapter 3. The YouTube section will also include a detailed introduction on its translation tool. The section on fan translations will give an insight into the origins of amateur translation, as well as the problems related to the concept.

3 YOUTUBE AND FAN TRANSLATION

YouTube is a website that was founded in 2005 and acquired by Google in 2006. Anyone with a YouTube or Google account can upload videos, comment on them or create playlists with their own or other people's videos. The user interface and general appearance have undergone multiple tweaks and changes, some features of the site have come and gone, but the essential function of uploading videos has stayed the same as it was a little over a decade ago. From a research point of view, the website offers multiple interesting functions: YouTube can be analysed as a technical tool, as a tool of video uploading and sharing, but it can also be analysed as a way of social influencing and information sharing. Section 3.1 will discuss YouTube from both standpoints, and section 3.2 will focus on its translation tool and introduce its user interface as well as its different functions in detail.

The tradition of fan translations relies heavily on the audience members' role as fans, their passion for whichever cultural product they admire or like, and their desire to help spread the product in question to members of their own linguistic groups, be it a non-mainstream web comic, an animation, or a fan fiction. The key theme in these translations is that they are made voluntarily by fans, for other fans (Pänkäläinen 2014: 1–2). In section 3.3, I will introduce fan translation as a phenomenon, its origins and development and some notable features, as well as its relation to professional translation and YouTube subtitling.

3.1 YouTube

Depending on the analyser's point of view, YouTube can be considered a beneficiary or a harmful platform. On one hand, it represents new media and content, technical development and the force of the Internet as a way to spread information. On the other hand, it is another competitor in the field of media, negatively affecting old business models and society, as well as being yet another space for online bullying. (Burgess & Green 2010: 15) From a more objective point of view, YouTube has its virtues and it is

extremely useful, but it also has flaws and problems, both as a tool and an online environment. In addition, a casual user views it differently from a content creator who earns their living making videos on YouTube, or a film producer who sees it as a threat to the industry.

In academic studies, YouTube has been described as "a site of online participatory culture" (e.g. Androutsopoulos 2013: 47; Burgess & Green 2009: 10). Jenkins et al. describe participatory culture as having five defining attributes: firstly, the barriers to artistic expression and civic engagement are low, meaning that creativity and social commentary via art is not discouraged. Secondly, sharing those products of creativity is supported. Thirdly, there is an informal power structure present, meaning that the most experienced members of that culture educate the newcomers. Fourthly, the members of that culture experience validation and feel that what they do matters. Finally, in a participatory culture the members feel at least somewhat connected to each other and give value to other members' opinions on their own creations. (Jenkins et al. 2007: 24)

Participatory culture is culture adapting to new media technologies, but it is also culture encouraging everyone to interact with media content (Jenkins et al. 2007: 25). YouTube offers its users exactly that: culture, art and social commentary in an online environment, older content creators are admired and copied, members interact with each other, either via videos or comments on those videos. YouTube's online culture is a unique space where "traditional culture", nationalities and cultural differences are all mixed together. Of course, it is heavily controlled and influenced by the Anglo-American world view as the original and main language of YouTube is English, but every content creator and every viewer also bring a part of their own culture and identity with themselves. YouTube is less defined by nations or nationalities and more defined by the individual people interacting in that space.

As is the current trend with websites, YouTube also adapts to its user's preferences and interests, suggesting videos that might be of interest based on the user's physical location or their previously watched videos. For example, a Finnish viewer who has watched motor sports videos in English might receive recommendations on rally race compilations

in English, Formula 1 videos on Kimi Räikkönen's radio conversations with the team, or the latest, most popular videos made by Finnish YouTubers. Everyone has a different user experience, carefully tailored by Google's algorithms, and it is therefore difficult to define a "typical" YouTube viewer or viewing experience. So, even though it is a site of participatory culture, everyone has a unique user experience.

From a more general viewpoint, the style and content of the videos are not restricted or governed by YouTube, but there are Community Guidelines which forbid certain types of content. Videos that violate these guidelines can be removed or at least demonetised, which means that the content creator will not have advertisements played during the video and so will not receive any advertisement revenues. The following list of forbidden content is paraphrased based on the official guidelines:

- Nudity and sexual content,
- Harmful or dangerous content,
- Hateful content, includes racist content,
- Violence and gore; excludes violence in a documentary, but it has to be explicitly stated,
- Harassing and bullying content, threats,
- Spam, misleading metadata; includes 'fake news',
- Copyright infringement, impersonation, and content endangering someone's privacy.

(YouTube Community Guidelines 2018)

As Michael Strangelove (2010: 4) puts it, YouTube is a social place instead of an archive. The main goal is not to collect and maintain knowledge and information: it is to let people upload any videos they want, provided that they do not contradict the Community Guidelines. It is a creative forum that can be used as a platform for social or political dialogue, sharing experiences, telling stories. In short, it is basic human communication in an online environment. YouTube describes its mission as "to give everyone a voice and to show them the world", and their values are based on the freedoms of expression, information, opportunity and belonging (YouTube 2018).

However, from another point of view, YouTube can be described as an archive, although an accidental one and not maintained as one (Burgess & Green 2009: 87). It is like a

library where the books are written by anyone, they can be read by anyone, it can be shaped by the browsers' preferences and interests, but there is no librarian to maintain the space or preserve its contents. That is to say: YouTube does not save the videos or upload backups to another server or a cloud service. If a video is deleted, either by YouTube or its original uploader, it cannot be retrieved by anyone else, unless a third party has illegally downloaded and re-uploaded it. Even then, it is not the same video with the same views, likes or comments.

From a marketing and business point of view, YouTube is an excellent platform to use targeted advertisement by utilising the users' search histories and browsing preferences. It should be effective and easy, considering that YouTube is now owned by Google and the majority of Internet users use Google Chrome as their preferred Internet browser. With five billion videos being watched daily (Coles 2018: 107), the coverage and amount of potential advertisements on YouTube is almost unbelievably large. Content creators can generally decide the amount of advertisements that can be included in their videos in the beginning, middle and/or end. Viewers have the possibility to skip the advertisements, apart from the first five seconds, which forces the viewer to acknowledge its contents which, again, is an effective way for marketers to gain exposure. The length of different advertisements varies, but some have decided to make the most of their five seconds to make their advertisements exactly that long, to effectively use the time most viewers are willing to grant them.

Consumers of mass media and advertisers are themselves turning into producers of online content (Strangelove 2010: 6). Every YouTube user is a potential competitor to content producers in the 'traditional media' and, looking at the number of subscribers that the most popular YouTubers have the ability to amass, the threat to conventional media producers is real. In order to compete for the same audiences, the crucial components to have besides content are quickness and approachability. Depending on the amount of editing in the videos, a YouTube vlogger can easily upload new videos even daily, and since they market themselves with their own faces, it is easier for a viewer to relate to a vlogger and his or her content than to that produced by a huge company. Vloggers also

have the ability to adapt more quickly to current events, trends, or the inside jokes of the Internet.

YouTubers would not exist without their audiences and YouTube itself would undoubtedly not exist without the audiences: in 2009, it is estimated that it cost Google 710 million dollars per year to operate YouTube without any profit from it (Strangelove 2010: 6). Almost a decade later, in 2018, it is difficult to access official data on YouTube's profits. Some sources estimate that it was 4 billion dollars in 2014 (Business Insider 2015), others suggest that it could be 9 billion dollars in 2016 and 13 billion dollars in 2017 (Investor's Business Daily 2016). Hiding the numbers could suggest that Google's acquiring of YouTube has not been quite what they had expected profit-wise. Profits aside, YouTube as a site is constantly developed and its features are updated: as mentioned in the Introduction, YouTube has over a billion users, and abandoning such a large group of users would be poor marketing.

Interaction between content creators and their audiences is one of the most notable features of YouTube. There are several ways in which the audience can participate in the experience and express their opinions. They might press the like or dislike buttons to quickly show their opinion on the video, and if they really like it, they might add it to their own playlist to be watched again later. They might share the video on social media – a feature that was removed a few years ago allowed people to post video responses that were shown in the vicinity of the original video (Strangelove 2010: 13). At the moment, the most effective way to interact with the content creators and their videos is to leave a comment in the comment section below the video. Complete anonymity is not possible, as the commenter needs either a YouTube or a Google account in order to post a comment.

A more concrete way for the audience to participate in YouTube video-making is to provide subtitles in different languages. It is a chance to visibly contribute to the content, and it can happen completely unprompted, provided that the content creator has opened the platform for subtitling the videos. Sometimes, the content creators might openly ask viewers to add subtitles: this is also known as modern-day crowdsourcing (Anastasiou & Gupta 2011: 637). A more subtle way to ask for subtitles is to use pop-up messages that

appear during the video. These pop-up messages can be personalised to the content creator's needs and may include features such as questionnaires, links to other videos or completely different websites, or a link to the subtitling platform, usually accompanied by the text "Help translate this video in your language" or some other variant of it. The following section inspects the translation tool more closely.

3.2 Subtitles on YouTube – YouTube's translation tool

YouTube's translation tool is a simple, straightforward subtitling tool, with limited features for ease of access. The following section will introduce the subtitling tool and demonstrate its use and features through screenshots. The video used as an example is, at the time of writing (16.2.2018), Dave Cad's newest vlog titled *What do we think about Finnish people?* Currently, it does not have proper subtitles in either English or Finnish, but a machine-generated translation in English is available.

The subtitles on YouTube videos can be divided into three different groups based on their language and agency: translations from a source language to a target language, translations in the source language (also known as closed-captioning in the US) and machine-generated or "automatic" translations. So far, machine-generated translations are mainly available in English-speaking videos, but they can be utilised in the closed captions that are made by human translators. YouTube's translation tool allows the translator to use the machine translation as a base for the translation and reduce the time required for typing out the entire transcription. Machine translation and language detection has not been perfected yet, so there are bound to be mistakes in the machine translation that the human translator must edit.

Ultimately, the responsibility for a correct translation is on the human translator: not only must the translation be grammatically correct, it also must make sense semantically and contextually. No matter how advanced the language detection might be, a machine cannot be held accountable for the mistakes it makes, not until we live in a future where artificial intelligence is at least on the same level with human consciousness.

The translation tool works with two kinds of translations: machine translations and human translations. An English machine translation is a default translation unless a translation has been added. The user may choose to show the translations automatically or they can be turned on manually using the icons in the bottom right corner of the video (Picture 1). The first icon puts on the default translation and the second allows the user to browse the different translation options. The other two icons are used to define the size of the video on the screen. The option to add a translation can be accessed by clicking the three dots below the video that indicate more options (Picture 1).
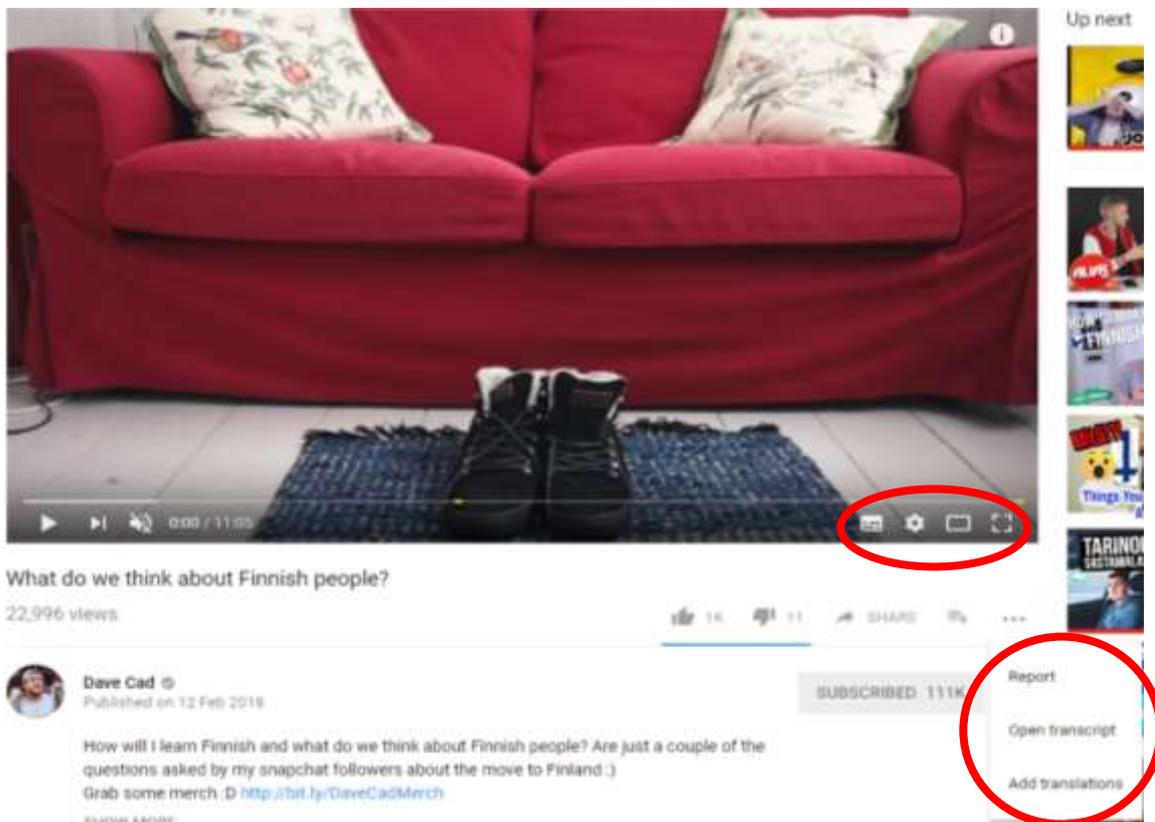

Picture 1. The basic layout of a YouTube video

Choosing "Open transcript" opens a text box next to or below the video, showing the transcript and time stamps (Picture 2). Since the video used in this demonstration at the time of writing does not have any translations yet, it uses automatically generated English

subtitles which usually have multiple mistakes: for example, the first line should start with "Oh hi everyone", not "Oh Haley one" as the machine has detected and suggested. The two phrases are pronounced similarly, especially if the speaker rushes the words, which is why it is easy to see how the machine might make the mistake. An inexperienced human translator might make the same mistake if they only rely on what they hear and do not take into account the context of the phrase.

Interestingly, the machine seems to recognise music playing in the video and indicates it using brackets. This is significant for any viewers with hearing impairments, as the use of music or silence can affect the overall feel or atmosphere of the video. Expressing the presence of music or any other sound is especially important in instances where it cannot be deduced from the video alone. For example, an on-screen explosion or orchestra performance does not necessarily need a verbal indication in subtitles, whereas an off-screen one does.
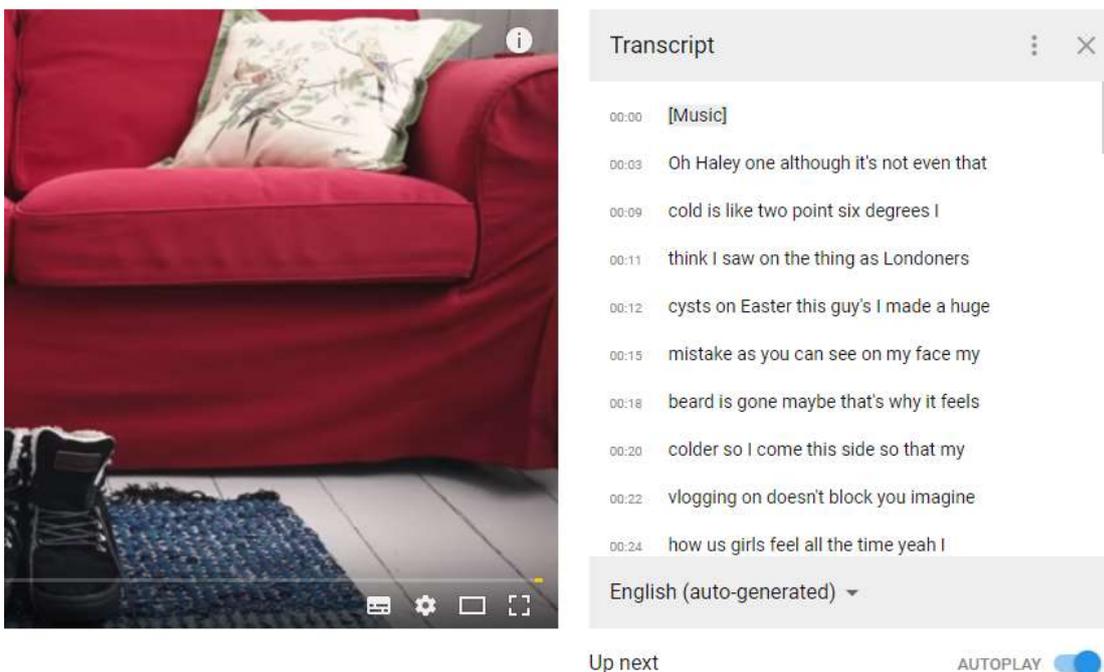
Picture 2. The transcript opens next to the video

Clicking "Add translations" takes the user to YouTube's "Creator Studio" (Picture 3), which first presents the option to add a translation for the video's title and description. The user also has the possibility to choose the language to which they wish to translate; YouTube gives automatic suggestions, in this case Finnish and English, but the user may also choose from 188 other language options. To make the process quicker, YouTube gives the option to use "Auto-translate" which produces a machine translation in seconds, after which the user can proofread it and edit the mistakes made by the machine. This study will focus on the subtitles, so the possible translations for titles and video descriptions will not be included in the material or analysis.
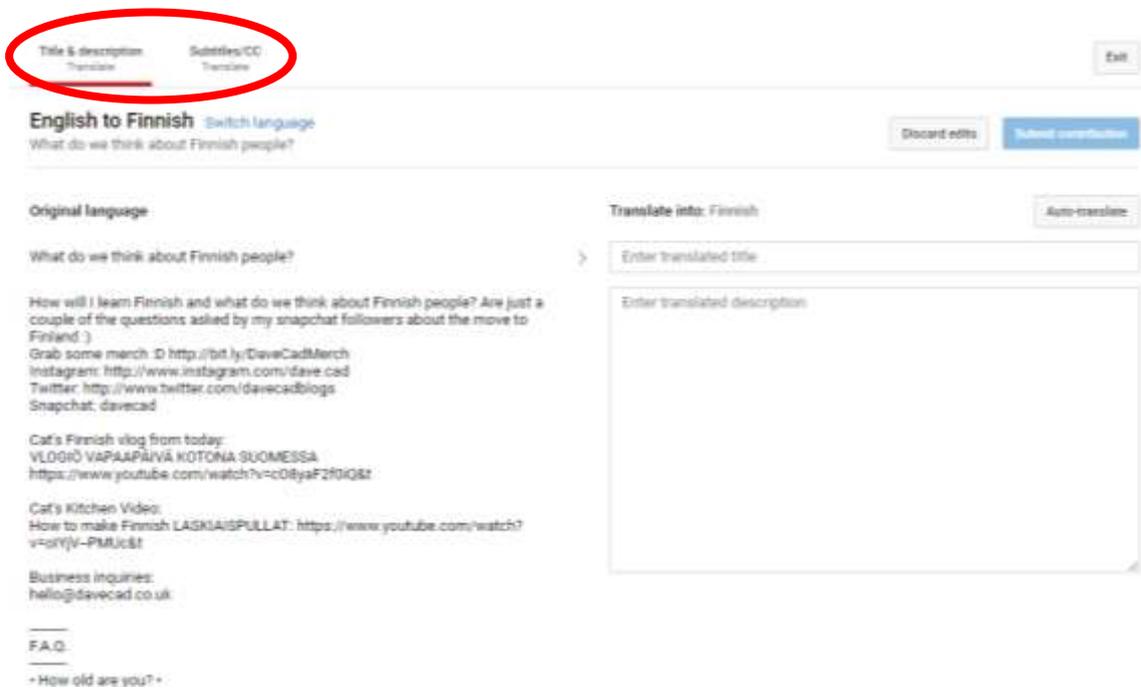
Picture 3. The original text in YouTube's "Creator Studio"

YouTube's translation tool automatically detects the timing of the speech, but the users can modify the time stamps to accommodate better for their translation solutions. From the "Actions" box (Picture 4), they may upload translation files or download their final products. The video that is being translated plays so the translators can see their process and product in real time, but they also have the option to pause the video while typing.

There is no character limit as to how long the subtitles can be, which means that it is possible to cover the whole screen with subtitles. Keeping this in mind, the length of the subtitles will also be discussed in the analysis of the quality of the subtitles in Dave Cad's videos in section 4.2.

The translation tool does not allow the user to modify the overall layout of the subtitles, i.e. they cannot be moved from the bottom centre of the video. This can be a problem, especially if the video incorporates text, or pictures, anywhere in the bottom third of the video. Either the subtitles will stay, covering information in the original video, or the timing of the subtitles is altered, which can lead to the violation of the principle of synchrony, a key feature of subtitling.
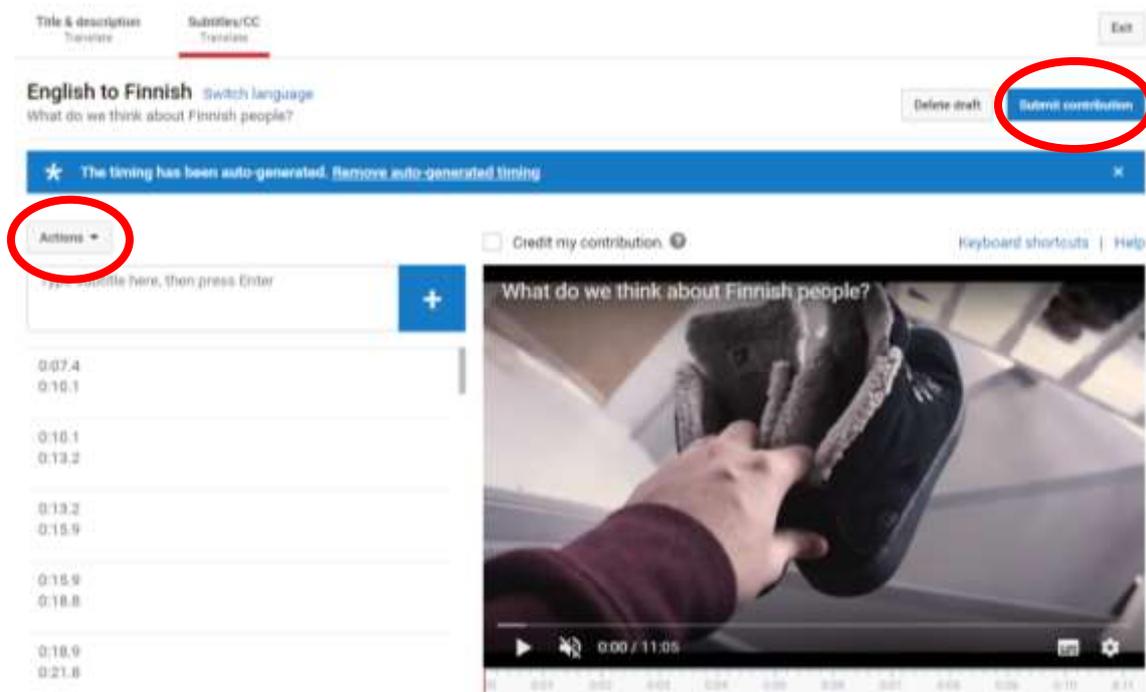
Picture 4. YouTube's translation interface

After finishing the subtitles, the translators also have the option to make themselves visible as the video's translator by clicking the option "Credit my contribution" as seen in Picture 4 above the video, and their YouTube username will show at the end of the

video's description. It is interesting that it is an option to remain anonymous as a translator when everyone else – including the commenters – must write with their name, or at least a pseudonym. On one hand, submitting an anonymous translation reduces the translator's responsibility for the quality of the subtitles, and anyone could write anything there, for instance racist comments without any personal consequences. On the other hand, it further diminishes the role, visibility and importance of the translator, even if that is the translator's own decision.

It is also possible to submit partly finished subtitles. When the translator has done as much as they can, they can send their subtitles by clicking "Submit contribution" (Picture 4). The subtitles can then be marked as finished, after which they will be checked and published, or unfinished, which leaves the file open for other translators who can continue and finish the subtitles.

Overall, it is clear that YouTube's translation tool is designed for amateurs: it is easy to use, plain, and many of the options are automated so that the only thing the translator needs to do is translate the speech and other possible text. This might be a conscious decision on YouTube's part to make the tool more accessible to everyone regardless of their technical know-how in order to get more videos translated, which in turn increases the sizes of potential audiences: larger audiences equal more ad revenues.

3.3 Fan translations

Fan translation as a concept is self-explanatory: it refers to translations made by fans who might not have any connection to translation as a profession. Fan translations and fan subtitles, or fansubs, have their origins in the distribution of Japanese animation, or animes, in Western cultures, and it has been a growing trend since the 1980s. The growing popularity of the Internet not only provided fans with tools for subtitling, but also the raw material, i.e. the anime episodes (Díaz Cintas & Muñoz Sanchez 2006: 37). One could argue that the popularity and demand for fansubs, before and even today, is due to the long waiting time for official translations. Even if an official English translation is to be

expected, it might take a couple of days at best until it is published, and fans are notoriously impatient.

The practice of fansubs raises some questions about copyright, and technically they are illegal. However, in the early days of the tradition, fansubs have been both acknowledged and approved by copyright holders. Fansubs have a potential to reach large audiences without the official parties' participation and, conventionally, when an official translation does get published or commercial distribution starts, the translators themselves discourage the use of fansubs. However, now that the Internet has made official distribution easier and anime has become more popular all around the world, the copyright holders' attitudes towards fan translators are changing. (Díaz Cintas & Muñoz Sanchez 2006: 44–45)

In addition to problems with the entertainment industry, there is another issue relating to fansubs or, more specifically, how fan translations affect the livelihood of professional translators. Fan translators might have some experience but they usually do not have any training in the area of translation and, as they make the translations as fans, more often than not they do it on a voluntary basis, meaning that they do it for free. (Pérez-González & Susam-Saraeva 2012: 151) From a professional and corporate point of view, the tradition of fansubbing is dubious, to say the least.

Examining the convention of YouTube subtitles in the light of these points of copyright issues and the capital losses of professional translators, it needs to be said that they do not apply in this case, specifically. Firstly, the whole idea of YouTube as a platform is that it is free to use, to both content creators and viewers alike, but there is a possibility to pay for the YouTube Red service which enables advertisement-free, offline and exclusive content (YouTube 2018). Secondly, since YouTube subtitling works on a voluntary basis, there is no money to be made on them in the first place. There are no losers in this equation, only winners: the content creators' potential audiences grow, and the translators get experience or just do it for fun. It could be argued, however, that this relationship between YouTubers and the translators is unbalanced, since the translators do not receive any compensation for their work.

Professional translators in Finland are fairly visible: in translated literature, their names are usually visibly written in the front covers of books, and in-house AV translators are mentioned at the end of films and programmes. Sometimes the AV translator is mentioned in connection with the translation studio that was commissioned to produce for example subtitles or voiceover. However, there are also occasions where only the translation studio is mentioned, and the individual translator is erased completely. Translators are traditionally given less attention than the creators of the original content, but nevertheless they have been given some credit for their work. It would seem that translators in the field of audiovisual translation are slowly turning into invisible translators, and not necessarily of their own will.

It is quite interesting that YouTube translators seem to be feeding this phenomenon of the invisible translator by consciously detaching themselves from their products, in this case the subtitles. What is more interesting is that, as YouTube translators are not paid in any way and they work on a voluntary basis, their only reward would be their moral rights to the subtitles, an acknowledgement of a work done and they are willing to reject it. Is it because they do not want to get acknowledged, are they indifferent about the work they have done, or are they afraid that they would be judged for their work or, rather, their mistakes? It is also possible that amateur translators do not consider their work as worth "owning" or as work in the first place.

The anonymity surrounding Internet culture and the use of pseudonyms offers the possibility to become invisible almost too easily. Human interaction is based on trust; that applies in an online context as well as in real life, but how can we trust something or someone we cannot see? People who need to rely on subtitles, for example because of a hearing impairment or a foreign language, place their trust in the translator to translate the content correctly so that they receive the correct information. If future translators grow up in an environment where their work is often unacknowledged, and they accept it as normal, will the invisible translator soon become more common than a visible one?

The following chapter will be dedicated to the case study where the focus will be on the translators and subtitles of Dave Cad's videos.

4 FAN TRANSLATORS AND THEIR SUBTITLES ON YOUTUBE

The aim of this thesis is to study the translators and their subtitles on Dave Cad's YouTube videos. The main interest lies in the quality of the subtitles, and specifically in the different types of errors that can be found in them. The study will consist of two parts and two sets of material; firstly, I will introduce the questionnaire and introduce as well as discuss the answers. The analysis will be mostly qualitative. Secondly, for the main goal of this study, I will analyse the quality of the Finnish subtitles in 17 videos that form the main material for the study. The focus will be on the translation errors found in the subtitles.

YouTube translations do not only differ from professional translations: they are also different from the "traditional" forms of fan translation as discussed in section 3.3. YouTube translation is legal, as long as the video itself is; it is easily available to viewers and translators alike; the platform offers a free, easy-to-use tool for subtitling. The activity is even encouraged: YouTube offers the possibility, and content creators ask for translations. It allows translators and language-savvy people to engage and create in the participatory culture of YouTube.

As the translators of YouTube are presumably amateurs, it is probable that they are not aware of subtitling conventions, not to mention translation theories and strategies. That is why it is interesting to see what amateurs consider as good translation, or at least good enough to publish. The point of translation error analysis in this study is to find what types of translation errors, if any, can be described as typical for fan translators, not to criticise their work. The interest is purely academic, as the work of fan translators is near impossible to govern.

4.1 Questionnaire answers

The questionnaire was constructed to gather both background information and information on how the translators work, as well as how they view their work. The

questionnaire was conducted using Google Forms, and the link to it was sent directly to the YouTube users who had allowed their name to be visible as a translator of a video and had included some contact information in their channel description. As the questionnaire relies on anonymity, the answers cannot be linked to a translator or his or her translation: by making answering anonymous, the goal was to get answers that are as truthful as possible. In addition, even incomplete answers were verbally encouraged, as it was assumed that the open-ended questions might make the answerers lose interest and decide not to submit their answers. In total, the link to the questionnaire was sent to 5 users, of which 3 submitted their answers.

The questionnaire (Appendix 1) consists of three separate sections: the first part includes questions about the translators' background, such as education and age. The second part includes questions about the translators' language skills, such as their native language and possible translation experience. The final part of the questionnaire includes questions about the actual process of translating YouTube videos, namely their problems and solutions but also their reasons for translating these videos. The collective answers are compiled and presented in Appendix 3, with an English translation in Appendix 4.

One of the biggest issues that arose during the gathering process was the unanticipated difficulty in contacting the potential participants. Many of the users I had planned to send a message to do not have a channel for contacting them, and YouTube's own messaging system is practically non-existent. This is why the focus group shrunk to only a handful of people and the focus of the whole study had to shift. Another significant issue was in the expectations that I had towards the answers and the actual answers I received. The open-ended questions were formulated in a way that would minimise the researcher's effect on the answers, but it resulted in rather short answers. This was the case especially in the part where the questions were about translation problems and difficulties: if there had been an example of a problem in the question, it could have prompted the answerers to agree with that problem and distorted the results.

Based on this process and the amount of potential and actual participants, an interview instead of an online questionnaire might have provided more substantial results. It would

have allowed additional or more specific questions where needed, and the interview situation itself could have prompted more thorough answers. If the potential participants are easy to contact and readily available, a questionnaire can be more efficient; an interview would work when working with fewer participants, but those participants would have to have a higher level of commitment and interest towards the study.

While the following results of the questionnaire might not prove scientifically significant or extensive enough for deducting information, they do provide an interesting snapshot to these fan translators' different backgrounds. All the respondents were under the age of 25 and two of them were under 18. One of the under-18-year-olds attended comprehensive school while the other attended upper secondary school; the third respondent had graduated from upper secondary school and works as a photographer and/or video maker. Everyone put Finnish as their native language, but one also put English as a native language. Other language skills listed besides Finnish and English were German and different levels of Swedish. No one had studied translation and only one had previous translation experience.

It was assumed that the translators would be young, but it was surprising to learn that one was in comprehensive school which means that they can be anything from 7 to 15 years old, although presumably closer to 15 years old. Finnish being everyone's native language was not surprising; however, one bilingual respondent was unexpected. One of the earliest hypotheses regarding the questionnaire answers was that people with multilingual backgrounds would be more likely to submit fan translations because of their backgrounds. This assumption was based on my own bilingual background and previous experiences with YouTube translation. Obviously, this study can neither confirm nor deny the hypothesis since other factors can also affect a person's decision to engage in fan translation, but it is interesting that there was one bilingual respondent, even in such a small focus group.

Some questions in the YouTube-part of the questionnaire were on the translators' motivation. All the respondents answered that they chose to translate Dave Cad's videos for fun, and two of them wanted to try what it would be to translate them. One respondent

had not subscribed to Dave Cad's channel, which could mean that they are not an avid watcher. When asked how they chose the videos they wanted to translate, one answered it was based on the content of the video while another chose it because it was difficult or challenging. One respondent submitted their own answer where they wrote that they were planning on translating multiple videos but so far had only translated one video. None of the answers indicated that the subtitles were made to gain practical experience.

When asked about translation difficulties or problems, only one respondent mentioned examples of different problems: they had had problems with timing the subtitles according to jump cuts, as well as translating idiomatic expressions. When asked how they solved those issues, they answered that there was no solution to the timing problem, and that they had to think of a solution to the problem with idioms. As a researcher, I would have wished for more concrete problems and solutions, but I also understand that the respondents may not think of their work as analytically or from the same point of view as professionals. However, something can be deduced even from this single answer: one, that idioms cause problems that translators, professionals and amateurs alike, meet often and two, that amateur translators might not realise that the solution to timing issues could be omission or condensation. These are, of course, the problems of one amateur translator, but at the same time they are not that different from the problems that also professional translators have.

Finally, the last questions were on the translators' overall translation experience on YouTube. All the respondents had also translated other than Dave Cad's videos. One had translated their own videos from English to Finnish, and they had also made subtitles for their own videos in Finnish for the hard-of-hearing. Another had translated videos similar to Dave Cad's content, also from English to Finnish. A third respondent wrote in their answer that they had translated videos "in foreign languages" into Finnish without defining what those foreign languages were. These answers support the fact that English is, in fact, the new lingua franca in commerce as well as entertainment, but that there are still people who value their native Finnish and feel the want or need to provide their fellow viewers with their native subtitles.

4.2 Analysing subtitling errors in YouTube videos

As stated before, there are no actual academic rules or guidelines to creating good subtitles. Different commissioners, such as BBC and YLE, have their own recommendations and technical requirements, but the responsibility for creating good subtitles is on the translator. After all, subtitling is context-reliant work, and therefore imposing a single set of guidelines could quickly present issues. The same applies to error analysis: the best approach has to be defined by the source texts. The error categorisation in the following section is simple and straightforward, partly due to the length of the study and partly due to the fact that conducting translation quality analysis on this scale would have taken more time than was available.

In order to produce reliable research, there are two issues that need to be discussed. First, the actual categorisation of the different errors. There are multiple instances where an error could be considered either a grammatical or a linguistic error – elaborated more on later in the text – but a decision had to be made between the two to avoid double-entries. In these cases, I categorised the errors based on consistency: similar errors would go to similar categories. The other issue is also related to the categorisation issue, but it involves the translator, as well.

As mentioned before, the translators can take advantage of the translation tool's machine translations and other translation applications, such as Google Translate. It became evident that some of the subtitles were created with the help of those machine translators, and some errors seemed to be caused by the machine, not the human translator. In those instances, a question arose: am I evaluating human or machine translation? As the purpose of this study is to find errors in human-made translations, is it right to collect those errors that are clearly made by a machine? In the end, I decided to include those errors, as it is impossible to distinguish and assign every single error reliably to either a human or a machine. The human translator acts as a filter of a kind between the "raw translation" and the final subtitles – as well as the audience – and should therefore be able to acknowledge and correct the errors made by the machine.

In order to count and categorise the different types of errors, the subtitles were copied and pasted from YouTube's subtitle transcription onto a Word file. At this point, the different error types were given a colour code to help distinguish the error types from each other and visualise the amounts and frequencies of those error types. The comprehensive results are presented in table form in Appendix 5. In order to make the results more readable, the following tables within the text are condensed, which means that some categories have been merged. For example, conjugation errors for verbs and nouns are separate in the appendix, but for the sake of clarity, they are merged in the text.

The categorisation for different translation errors for this study is visible in Table 1. The error types are divided into four groups: grammatical, linguistic, punctuation and technical errors. The categories have been constructed based on Vehmas-Lehto's (2005) translation error analysis and Díaz Cintas & Remael's (2007) conventions on subtitling, as well as the Finnish conventions for subtitling. For grammatical errors, the guidelines of Kotus (Kotimaisten kielten keskus or Institute for the Languages of Finland) will be referred to as an authority on the Finnish grammar.

Table 1. Error categories

Error type	Amount
Grammatical error	199
Linguistic error	252
Punctuation error	1870
Technical error	1097
Total	3419

Grammatical errors are errors in the target language, such as wrong inflections or wrong tempus. They differ from linguistic errors in that, in this occasion, a linguistic error relates to the translation process and the translation of words and phrases, such as idioms. Punctuation errors are spelling mistakes or deviations in the use of commas, dots or capital letters. It is expected that especially errors in the use of capital letters are due to

interference from the English language. Table 1 above shows the distribution of errors across different categories. Quite interestingly, the amount of grammatical and linguistic errors – even combined – only accounts for a little more than 10 percent of the total amount of errors. On average, these numbers suggest that each subtitled video contains 11 grammatical and 14 linguistic errors, which is still substantial for 10-minute videos.

Over half of all the errors counted were punctuation errors, and about third of all the errors were of technical nature. Taking into account that the translators are amateurs and not aware of translation's technical conventions, the number of technical errors could have been even higher. On average, each subtitled video contains 110 punctuation and 64 technical errors. Error distribution across different videos has not been noted, as the goal of the study was not to rank the subtitles based on the amount of errors, but rather discover the types of errors made. Some of the subtitles had very few mistakes while others had some type of error on every single line of subtitle, but all of the subtitles did have errors.

Grammatical, linguistic and punctuation errors have been divided into separate categories, as they can be found on different 'levels' of language. During the analysis process, it became clear that there is one instance of a translation error that could potentially be categorised either as a grammatical or linguistic error but which I eventually decided to classify as a linguistic error. This instance was the English passive 'you' and its translation into an active 'you' in Finnish, which is by no means a new issue.

The English passive 'you' has often been mistranslated into Finnish as an active 'you', either as singular 'sinä' or plural 'te', even when no specific entities are being addressed. While it is widely used, specifically in spoken language, it has in this study been classified as a linguistic error, as the use of a passive voice would be more natural in the Finnish language. It is quite easy to see why the translators of YouTube videos would mistake the passive 'you' into an active one: in his videos, Dave Cad addresses his audience directly, as if talking with them.

This also relates to another translation issue: when Dave Cad uses an active 'you', as opposed to a more unambiguous 'you guys', does he mean an individual viewer or the

audience as a whole? There were a couple of instances in the source material where the Finnish singular 'sinä' was used when the plural 'te' would have been a more natural choice. A plural 'you' is the obvious choice in traditional media, such as TV and film, where the individual viewer is seldom addressed directly. However, given the more personal nature of YouTube vlogs, a singular 'you' could be a valid solution if it is used to solidify the personal aspect of the video. In the end, the correct choice depends on the context.

Finally, technical errors relate to the technical aspect which in turn is based on the Finnish subtitling conventions. The length of the YouTube subtitles will be compared to the subtitling conventions on Finnish television subtitles. This means that the subtitles should not be longer than 35 characters per line, nor should they be longer than two lines. As it is highly unlikely that the translators have had training in AV translation or subtitling or are aware of the technical limitations and aspects such as reading speed, translation errors relating to the abovementioned technical aspects of subtitling are highly likely. The translator may not be aware of these errors, or they might acknowledge them and choose to ignore them in order to, for example, stay true to the content of the videos.

All of the different error types were given a colour code to help distinguish the error types from each other and visualise the amounts and frequencies of those error types. The errors were marked in their respective colours on the subtitles that were copied onto a Word file, after which the error types were collected, listed, and counted. The comprehensive results are presented in Appendix 5. In order to make the results more readable, the following tables within the text are condensed, which means that some categories have been merged. For example, conjugation errors for verbs and nouns are separate in the appendix, but for the sake of clarity, they are merged in the text.

Different error types, namely grammatical and linguistic errors, are given examples, each of which have been chosen randomly from the subtitles. Each example includes a reference to Kotus's guidelines for the correct use of Finnish. The time stamp of the subtitle is also included. Any Finnish used in the actual text is written in italics instead of quotation marks to avoid mix-up. This includes my versions or suggestions for better

translation solutions. The titles for the different videos where the examples can be found are written as they are. This means that if the original title is written in all-capital letters, it is written the same way in the text, as well.

4.2.1 Grammatical errors

Grammatical errors are listed in Table 2. They relate to the correct use of Finnish grammar, namely conjugations. Conjugation errors have been divided into two different groups: unconjugated and falsely conjugated words. Both groups include verbs, nouns and adjectives and/or adverbials with conjugation issues. Table 2 shows the total sum of these errors, and the comprehensive list with each word category is available in Appendix 5. Wrong pronouns and wrong conjunctions have their own groups, as it did not seem appropriate to merge them with any other groups, and they both did have a substantial number of errors as they were. The group for 'invalid expressions' includes missing sentence elements, falsely used singular and plural expressions within a sentence, and colloquial expressions.

Table 2. Different types of grammatical errors

Grammatical error type	Amount
Unconjugated words	55
Wrong conjugation	66
Wrong pronoun	26
Wrong conjunction	20
Invalid expressions	22
Total	199

Colloquial expressions are listed as a grammatical error rather than a linguistic one even if it could technically fit into either category. In instances where both formal and colloquial Finnish were used, the erroneous expressions were determined based on the consistency of the language used: if colloquialisms were consistently used throughout the video, they can be considered a translation strategy and thus pass the error check.

However, if there were a few expressions of colloquial language in subtitles that were mainly written in formal Finnish, those instances counted as an error. Invalid expressions in Table 2 refer to sentences or expressions that do not work in Finnish, as they, for example, lack a component needed for a full sentence. Interference from English is assumed to at least partly contribute to these errors.

Perhaps surprisingly, grammatical errors form the smallest group of errors with the total amount of errors being 199. The amount of unconjugated words – most of which were nouns – is most likely affected by utilising machine translations. Since English uses very little conjugation compared to Finnish, and Finnish conjugation varies depending on other sentence elements, it is difficult for a machine to accurately identify the correct form. Unconjugated words might also have been left there from a raw translation that the translator has forgotten or decided to ignore from one reason or another. Conjugation issues amount to 121 grammatical errors, which is over half of all the errors in this category.

In example 1 from Dave Cad's video FINLAND & ENGLAND (The Differences), the predicative adjective *hiljainen* has been left unconjugated. In this instance, the correct form would be the form *hiljaista*, as the *-nen* suffix requires a partitive form (Kotus 2020h).

(1)   01:31 On todella hiljainen

Example 2 is from Dave Cad's video Am I A Feminist? (Snapchat Q&A!) These subtitles in particular had multiple errors, as it would seem that machine translation was utilised but not edited. Compared to other videos, these subtitles have a remarkable amount of grammatical errors, such as the unconjugated verb *nähdä*. This sentence would need the conditional *näkisin* in order to express the possibility of a desired action (Kotus 2020c).

(2)   02:34 Mielelläni nähdä niin paljon kuin mahdollista

The single largest group of grammatical errors with 66 entries was words that were conjugated falsely, either verbs, nouns, and adjectives and/or adverbials. Example 3 shows the verb *ehdotella* conjugated in the singular form *ehdotellut*, when the it should, in accordance with the plural subject, have the plural form *ehdotelleet* (Kotus 2020j). This example is from the video REACTING TO FINNISH MUSIC VIDEOS | Part 3.

(3)   00:08 Olette ehdotellut minulle uusia videoita

There were a few cases where wrong pronouns or conjunctions were used. Example 4 from the same subtitles as example 3 showcases the use of an incorrect pronoun, or rather, the use of a pronoun in an instance where it actually alters the meaning of the sentence. To add context, example 4 is from a point in the video where a video by the Finnish pop singer Robin is being reviewed, and he seems to be wearing a similar red jacket that Justin Bieber has worn in a video. The word *videossaan* in itself includes a possessive suffix and implies possession between the subject and object of the clause. However, adding the possessive pronoun *hänen* changes the ownership of the object (Kotus 2020e). The use of the pronoun *hänen* in this sentence implies that Bieber has worn a similar jacket in one of Robin's videos, when that is most definitely not the case.

(4)   03:22 Eikö Justin Bieber pidä päällään samanlaista punaista takkia yhdessä hänen videossaan?

Example 5 is from the same subtitles as examples 3 and 4. From my personal experience of proofreading other students' Finnish texts, relative clauses and the use of corrective relative pronouns *joka* and *mikä* and their variants is difficult to many people. In short, the former is used usually when the relative clause refers to a specific word or word pair, whereas the latter is used only when a whole sentence is referenced (Kotus 2020i). In this case, the subordinate clause should use the pronoun *joita* instead of *mitä*.

(5)   00:27 eli tänään minulla on kahdeksan musiikkivideota, mitä te olette ehdottaneet minulle

The last error category, invalid expressions, is the most mismatched group, as it includes error types that did not group well with any of the other groups. Example 6 from the subtitles of the video DRONE PROBLEMS shows the use of the colloquial form of *näämme*, instead of the formal *näemme*. As elaborated earlier, a colloquialism was considered an error if they were not used consistently throughout the video, and in this case, the colloquial style was not consistent even within a single sentence. A more acceptable version would have been, for example, *En tiiä, näämme sen kun päästään sinne.*

(6)   07:06 En tiedä, näämme sen kun pääsemme sinne.

There were some instances where whole sentence elements were missing from the subtitles. Example 7 includes also the previous sentence, as it is important for understanding what exactly the sentence is missing. These subtitles are from the video titled 2017 [FIN/ENG Subs]. The second subtitle is actually missing two elements; firstly, there is no verb, and secondly, as a relative subordinate clause, it needs a relative pronoun. Correcting the other errors in the sentence, it should read *joka on suomalainen beatboksaaja.*

(7)   02:35 Tein myös tämän videon Felix Zengerin kanssa,
        02:38 suomalinen beatboxaaja.

Finally, there were a few cases where singular and plural forms were used erroneously, such as in example 8 found in the subtitles for the video REACTING TO FINNISH MUSIC VIDEOS | Part 7. In this instance, the plural noun would require the plural form *vuotavat* for the verb. In spoken language this passive expression is common, but as the rest of the subtitles were mostly written in formal Finnish, it counted as an error. As such, it could have also been categorised as a use of colloquialism, but it is less notably colloquial language than, for example, the use of *sä* or *mä*.

(8)   03:53 Korvani vuotaa verta

Grammatical errors were possibly the most difficult ones to collect and categorise as there were multiple cases where there was clearly a mistake, but it was hard to explain how exactly it was a mistake. There must undoubtedly be better, more sophisticated ways for analysing grammatical errors, and my method was crude, at best, but it was interesting to be able to present them in a numerical form. Compared to the other categories, grammatical errors formed the smallest error category, but based on these figures alone it is impossible to say if it is a positive outcome. Similar research could be conducted on different subtitles made by translators with varying levels of skill, education, and work experience. Those results could be cross-referenced with each other to determine how good or bad fan-made subtitles on this platform really are.

4.2.2 Linguistic errors

Table 3 lists different types of linguistic errors, namely mistranslations. Here, a linguistic error is a falsely translated idiom, a falsely translated word or word chain, or words or word chains that have been left untranslated. A direct word-for-word translation that does not exist or work in Finnish has also been counted as a linguistic error. This means that passages that seem to be translated by a machine are regarded as a linguistic error, as the human translator has not decided to modify it to read more naturally. The category for the false translations of 'you' include the aforementioned passive voice as an active and the use of a singular 'you' instead of a plural.

Table 3. Different types of linguistic errors

Linguistic error type	Amount
Idiom translated falsely	10
Words/sentences translated falsely	182
Words not translated	58
'You' translated falsely	2
Total	252

As with grammatical errors, some types of linguistic errors have been merged to form a single group, and a more detailed error list can be found in Appendix 5. The group for words/sentences translated falsely also entails the word 'it' translated directly in occasions where it should not be translated, word-for-word translations, and the use of a different person in the translation which also includes the use of a passive form instead of an active form, specifically in cases where the subject is clearly stated. Words not translated also includes untranslated interjections.

Again, compared to the other categories, linguistic errors formed a surprisingly small group of errors, with slightly more cases than grammatical errors. Even the amount of idiomatic errors was surprisingly lower than expected, with only 10 cases. Example 9 from the video FINLAND & ENGLAND (The Differences) shows the English phrase 'freeze your butt off' translated literally, even though it is impossible to freeze something off in Finnish. In this case, the best solution would be to translate the meaning, which is that the weather is extremely cold, with a sentence such as *Talvella on törkeän kylmä*.

(9)   02:36 Talvella takapuolesi jäätyy irti

Plain false translations were the single largest group within linguistic errors, amounting to over half of them. In some cases, such as example 10 from the same video as the previous example, the translator has not understood the meaning of the original word or deduced it from the context. This example includes the whole sentence to show the context of the erroneous translation. The original sentence in English describes the British seasons and weather, mentioning that they have 'this kind of grey wash over everything.' This would best translate as a mood or a feeling, as it is an abstract concept or perception of one's environment and has nothing to do with the actual concept of washing. In fact, the easiest way to solve this would be to remove the word *vesipesu* entirely and use the adjective *harmaa* to describe the weather more accurately.

(10)   02:18 Suurimman osan ajasta meillä on vain sellainen
        02:19 harmaa
        02:20 vesipesu
        02:21 kaikkialla.

Example 11 from the video My Finnish Summer Begins! showcases a situation where two similar English expressions have been confused with each other and the wrong one has been chosen for the translation. In this case, the English expressions in question are 'fixed on you' and 'fixated on you', the former of which is used in the video in the context of a camera being fixed on someone. However, the translator seems to have interpreted for it to mean something more similar to the latter expression, suggesting an emotional attachment rather than a physical one. A more natural way to express this in Finnish would be to use a simpler verb and translate the sentence as something along the lines of *Se kuvaa juuri sinua*.

(11)   03:06 Se on kiintynyt sinuun tai jotain!

Another reason for errors caused by misunderstandings or assumptions are homophones, words with different meanings that are written differently but pronounced similarly. Example 12 from the video 100,000 SUBSCRIBERS Q&A shows the mix-up between the words 'rod' and 'road'. The topic of conversation is clearly fishing, so the reason for this error must be misinterpretation or the lack of knowledge of the word 'rod', otherwise they should have been able to translate it correctly as *ottaa onkesi mukaan*.

(12)   06:43 Minulle on mahtavaa se, että voit melkolailla mennä kaikkialle ja ottaa tiesi mukaan

Untranslated words or sentences are left untranslated perhaps because the translator could not solve an issue or did not recognise the meaning of the word or sentence and decided to submit an unfinished translation out of lack of interest or time. It could also have been left there by mistake, for example if it was skipped during the original translation process and then simply forgotten in the final product. Machine translators in particular seem to leave portions they do not recognise untranslated, especially if the original text is a play on words. This can lead to such gems as in example 13, from the subtitles in Dave Cad's video Am I A Feminist (Snapchat Q&A), where the machine has translated the only word it has recognised in the expression 'Abso-bloody-lutely':

(13)   03:01 ABSO-verinen-lutly!

Worth mentioning are also the two instances where a clearly passive 'you' has been translated as active. The amount could have been and actually was expected to be higher, as Dave Cad uses it often in his speech, which is why it was originally given its own error category. A passive *sinä* can be used in colloquial language or as a rhetorical device (Kotus 2020f), but in itself *sinä* refers to second singular person. In example 14 from the video 100,000 SUBSCRIBERS Q&A, Dave Cad has just talked about how much he loves the density of Finnish forests and how 'you can just walk and walk'. The use of *sinä* in this case creates a clumsy sentence which would work much better if it was reformatted completely, for example as *Siellä voi vain kävellä ja kävellä.*

(14)   06:39 Sinä vain kävelet ja kävelet ja kävelet

Considering the linguistic aspect of translation, some notions arose during the gathering process. Namely, that there were several issues that, while not necessarily plain errors, affected the readability of the subtitles. For example, in some cases translating every single word actually produced a sentence that sounded very strange in Finnish. Condensing and omitting are key elements in subtitling which sometimes cause problems for the translator, but sometimes they absolutely should be utilised. Oftentimes, filler words such as 'like' have been translated even when they do not provide any additional information.

On word level, one issue that arose was with the use of gendered nouns. In English, 'dude' has become more gender neutral and it is used more flexibly. However, translating it into Finnish as *jätkä* focuses it more towards the male audience and sounds odd. A gender-neutral *tyyppi* would avoid dividing the audience based on their gender. Another word-level issue was the literal translation of the words 'it' and 'it is' as *se* or *se on* in occasions where it does not sound natural in Finnish and should be replaced with a passive form or avoided in some other way.

Finally, and this is also partially a punctuation issue, in some cases repetition was emphasised with capital letters. In Internet speech, capital letters are traditionally used when indicating loud noises, e.g. shouting or explosions, and by now every Internet user

should know that. Repetition could more easily be emphasised by using bolding or italics, both of which can be easily done even in YouTube's own translation tool.

4.2.3 Punctuation errors

Deviations from the Finnish punctuation conventions are listed in Table 4. They are divided into errors in the use of capital letters and punctuation marks, such as commas and dots. Spelling mistakes or typos are also regarded as a punctuation error rather than a grammatical error. One way to distinguish a punctuation error from a grammatical error is to check the keyboard layout: if it is clear that there has been a mix-up between two adjacent keys, the error can be classified as a punctuation error. The use of comma was determined based on Kotus' guide (2020g): two main clauses that do not share a sentence element are separated with a comma, as are relative and subordinate clauses. Interjections in front of a sentence are also usually separated from the main clause by a comma. However, in cases where there are two main clauses and the other is significantly shorter, meaning 4 words or less, the comma can be excluded in favour of readability.

Table 4. Different types of punctuation errors

Punctuation error type	Amount
Capital letters	335
Commas, dots, etc.	1307
Typos/spelling mistakes	228
Total	1870

Capital letter errors include missing capital letters at the beginning of a sentence, missing capital letters in proper nouns, as well as the use of capital letters in places where it should not be used. Example 15 from the video 2017 [FIN/ENG Subs] includes two instances where a capital letter is used unnecessarily. Firstly, as the second line of subtitles should end in either a comma or a colon, the third line of subtitles is not a separate clause and

should not have a capital letter. Secondly, as opposed to English, Finnish does not capitalise languages. Therefore, *ruotsissa* should not have a capital letter.

    (15)   03:47 Joten 2017 vuoden lopussa
                 03:49 ja sanon sen tässä
                 03:50 Minä todella haluan olla jotenkin hyvä Ruotsissa.

The category for errors in punctuation marks was the single largest error category in the entire study, with missing dots overrepresented with 750 instances. Other types punctuation mark errors were missing commas between main clauses, before a relative clause, or before a subordinate clause, additional punctuation marks, missing dots, exclamation or question marks or other missing punctuation marks, incomplete ellipses, ellipses with too many dots, missing quotation marks, or the use of an asterix.

The vast amount of punctuation errors seems to suggest that, in this relatively new medium, old writing guidelines regarding punctuation are easily discarded. It is interesting, considering that the purpose of punctuation marks is to parse sentence elements in order to make reading and comprehending the text quicker and easier. This is especially true for subtitles, as they are rarely re-read multiple times and should be understood immediately.

The examples I have chosen for punctuation mark errors are those that have a more exciting explanation as to why they are errors. Example 16 from the video MY VISIT TO TANZANIA (Serengeti & Zanzibar) shows a missing hyphen between the proper noun Ngorogoro and the following word. In Finnish, compound words with a proper noun as the first part need a hyphen between the words (Kotus 2020k). Therefore, the correct form is *Ngorogoro-nimisen.*

    (16)   01:51 Juuri nyt olemme Ngorogoro nimisen kraatterin reunalla.

Example 17 from the video REACTING TO FINNISH MUSIC VIDEOS | Part 4 shows an incorrect ellipsis. An ellipsis indicates an incomplete or interrupted sentence or clause

with three dots (Kotus 2020d). Example 17 has both an incorrect ellipsis with four dots. as well as an additional space between the word and punctuation marks.

      (17)   03:32 Ja he vain .....

Numeral errors are errors in expressing numbers and numerical values. According to international standards, time is expressed as can be seen in the time stamps in these examples: the hours and minutes are separated with a colon. However, in Finnish, they should be separated with a dot (Kotus 2020a). Therefore, in example 18 from the video My Finnish Summer Begins! the time of day should be written *6.40*.

      (18)   00:01 Kello on 6:40 aamulla.

The final example in punctuation mark errors is the use of asterixis. It has different uses, most notably as a break in search engine searches, as a means to indicate a non-grammatical form in linguistics, and as a means to indicate a clarification in a footnote (Kotus 2020b). However, in some cases, such as in example 19 from the video where Dave Cad is apparently doing something silly, the asterixis in the subtitles are used to indicate the action happening on-screen.

      (19)   04:34 **hölmöilee hetken...*

It is difficult to pinpoint a single reason for the different punctuation errors. Capital letter errors, while in some cases are due to interference from English, are also likely due to negligence and disregard towards writing guidelines. Starting a clause with a capital letter and ending it in a dot is one of the first things taught at school, so ignorance should not be the cause, or an excuse. Errors in the use of a comma are a little more understandable, as there are multiple cases where it is up to the writer to decide whether or not to use a comma, but even then there are cases where a comma has to be used, if only to make reading easier. One possible way to investigate different causes would be to collect a target group of amateur translators and have them complete a simple test on punctuation.

4.2.4 Technical errors

Finally, technical errors found in the subtitles are listed in Table 5. A technical error is a line that has over 35 characters, or a timing error which includes subtitles that stay on-screen for too long or start too late. The category "other technical errors" entail subtitles where the line break is falsely constructed and subtitles that are too short (e.g. one word per subtitle instead of whole sentences). One-word subtitles are not counted as an error, however, if it can be said that it is used for emphasis or it is suited for the context of the original video.

As expected, the most common error was the length of the subtitles on a single line. It is possible to divide the text on multiple lines in YouTube's translation tool, but as it does not count the characters, calculating and dividing the characters would increase the translator's workload. Thus, it is understandable that the translator would prioritise the content of the subtitles over its appearance, also considering that the translator is not even likely to be aware of reading speed recommendations or character limitations in professional subtitling.

Table 5. Different types of technical errors

Technical error type	Amount
Too many characters per line	980
Timing errors	57
Other technical errors	60
Total	1097

YouTube videos are generally cut differently: they are usually more fast-paced to produce more efficient and compact videos, as opposed to "traditional" video making where the editor tries to make the video progress as smooth as possible. This also presents an issue with subtitling where, traditionally, a set of subtitles is timed with cuts to ensure a smooth viewing experience. Sometimes YouTube videos are cut in a way that leaves very short

"scenes" where only a word or two is uttered. Sometimes multiple similar "scenes" with only a word per cut are put together, producing a disrupted sentence.

From a translator's point of view, the resulting subtitles are somewhat unorthodox: instead of stringing the words together into a coherent sentence that stays on-screen for the duration of that sentence disregarding the cuts, the subtitles often follow the original audio loyally, dividing the sentence according to the cuts. The other extreme is long subtitles that stay on-screen for too long. They usually result from a longer scene in the video that has not been cut mid-sentence, and the translator has not decided to divide it into multiple subtitles, either because of laziness or ignorance. This study material for example included subtitles that stayed on-screen for eight, even ten seconds.

Example 20 from the video LEARNING FINNISH | PART 1 shows some of the technical errors: the first line of subtitles is too long, whereas the other subtitles are too short. The subtitles *"aasi"* and *"sorsa"* are necessary for the deaf or hard-of-hearing, but they could have easily been combined into a single-line subtitle. There are also issues with punctuation marks: some commas and dots are missing. In addition, the words *lammas* and *kissa* in the first subtitles should be conjugated, or other words should be added to accommodate the unconjugated ones.

> (20)   01:54 joten siksi sitä kutsutaan "lammas" ja tiedän "kissa" joten viimeinen
>   on apina
>   02:01 "aasi"
>   02:02 Se on aasi
>   02:04 "sorsa"

Some of the issues with subtitle length are due to YouTube's automatic subtitle allocation which separates the sentences based on pauses in speech. However, automatically generated timing can be disabled, which means that the timing of the subtitles can be toggled by the translator to generate longer and more "traditional" subtitles. It is possible that, as amateur translators rarely have academic knowledge or real-life experience in translation and subtitling, they are not aware of the fact that the subtitles would be easier to read and view if they were in a more coherent form. Adjusting timing can also be seen

as too time-consuming and nit-picking, suggesting that amateur translators possibly have a "close enough" approach to translation.

The same can be said to apply to the amount of characters in the subtitles. During the analysis process, it became evident that lengthy subtitles were an issue. There were also multiple cases where the subtitles could have been divided on two lines to avoid too long subtitles but were instead written on one line. On one hand, the viewer can pause the video in order to read the subtitles, but on the other hand it also disrupts the viewing experience and causes extra work for the viewer. YouTube's subtitling tool does not count the characters, it does not limit the amount of characters, and thus it does not suggest an optimal length for the subtitles, as opposed to many other subtitling tools used by professionals. Even if YouTube's tool is for amateurs, it could feature a suggested character limit with an explanation to the translators.

5 CONCLUSIONS

Society's success is often measured by its technological achievements: development is crucial for the preservation of culture and, indeed, humanity as a whole. Our current technology is advancing in all area of expertise, but especially in information technology. The Internet has enabled developments and discoveries and lead us to a point where practically the whole world is at our feet, or rather, on our screens. We see and hear people and their thoughts which would otherwise go undiscovered, we consume media products on different platforms, and we do it all across language barriers.

Audiovisual translation has developed and evolved following the development of technology and different media. From film translation to subtitles on streaming services to localised website translation, new techniques, tools as well as professions have emerged. The field of audiovisual translation is constantly in the middle of turmoil and change, making it a never-quenching source and subject of research. Changing the point of view from technology to humanities, a whole new selection of research questions are available: how does it all affect people, their roles and working habits, how can these advancements help or hinder certain demographic groups, are all of these developments ethical?

The main purpose of this study was to study amateur or fan translators and their subtitles in the context of YouTube, a video sharing platform. The focus was specifically on the different types of errors that could be found in the Finnish subtitles of Dave Cad's videos. Dave Cad is a British YouTuber whose content is heavily influenced by Finland, which is a rare combination, but which also made for an interesting source of material. Quite disappointingly, due to the poor answer rate of the questionnaire participants, I only received answers from three translators. The number of translators that were contacted was small to begin with, which is why the focus point of the study had to shift from participants to products.

The aim of the questionnaire was to collect information on the translators' backgrounds: their age, education level, professions, and language skills. In addition, they were

questioned on why they decided to translate YouTube videos, as well as the possible problems they faced. The results were hoped to reveal some common factors between the participant that could have indicated an inclination towards translation as a profession, either through their education or interest towards the field. The consensus based on this questionnaire seems to be that YouTube videos are translated more just for fun and less for experience. It is impossible to make any generalisations based on the results gathered from these answers: it is also difficult to say whether or not a larger research group would have provided different results.

The translators were also asked if they had any issues or problems whilst translating the subtitles. Again, the answers provided little insight into the minds of the translators with answers such as 'no' and 'some, but I had to come up with something'. It is possible that the participants did not fully understand the question or did not recognise the concept of a translation problem or issue. I decided to not include any examples of errors in the questionnaire in fear of affecting the results. During the error analysis, it became evident that the participants should have encountered at least some issues, even if with just words they did not know how to translate.

After the re-shift in research questions and methods, the subtitles and errors therein became the primary material for this study. Translation quality assessment was conducted on subtitles in 17 YouTube videos by Dave Cad, and the material amounted to approximately 15,200 words. The subtitles were assessed based on four error categories: grammatical, linguistic, punctuation, and technical errors. The categories were mainly inspired by Vehmas-Lehto's error analysis model, but they were simplified significantly, and even then there were in total 47 different error types that could be found in the subtitles. The different error types have been condensed and combined into 15 categories in the actual text, for example combining unconjugated verbs and nouns into the same category.

Grammatical errors, in total 199 separate instances, were quite surprisingly the smallest error category. Most of them were conjugating errors, and even some of those were most likely made by the machine translator used by YouTube's translation tool. As it would

have been impossible to accurately distinguish between errors made by humans or machines, every error was counted. After all, the human translator should have at the very least edited the machine translation to comply with Finnish grammar standards.

Linguistic errors were the second smallest error category with 252 instances in total. Linguistic errors were issues with the actual translation, as opposed to grammatical errors that were issues with the use of the Finnish language. Mostly, linguistic errors were words or expressions that had been translated falsely, either because of a misunderstanding or a misjudgement of context. Words or sentences left untranslated also formed a large part of the total amount of errors, which can also be explained by the use of machine translations. As machine translators have difficulties distinguishing between the correct Finnish word forms and the original English words, the less obvious cases are left untranslated.

The category for punctuation errors see an influx in numbers, as the total amount of punctuation errors is 1870. The results in punctuation errors are skewed by the astonishing number of missing dots, mostly at the ends of clauses. The missing dots alone amounted to 750 errors, overshadowing the second largest group of errors in the use of capital letters which amounted to 335 errors. It is assumed that most punctuation errors can be found in the subtitles because the amateur translators did not deem them necessary. Based on analysis alone, it is difficult to present accurate causes for the errors in grammar, translation, or punctuation. Interference can be the cause in some cases, namely in the use of capital letters and in idiomatic expressions. The origin of punctuation errors must be somewhere else, possibly in the way Internet changes the traditional writing norms, which could prove an interesting topic of research.

The category for technical errors was expected to register multiple errors, and it was the second largest error group with 1097 issues in total. This was expected, based on the translators' generally low level of knowledge on subtitling conventions and recommendations. Most of the technical errors, 980 of the total number, were subtitles that exceeded the recommended 35 characters-per-line format. YouTube does not take responsibility for the quality of subtitles on its platform, just as it does not take

responsibility for the quality of the videos uploaded to the platform, only acting on cases that clearly contradict its guidelines. However, there are no guidelines for subtitles, although it is clear that the platform could benefit from them. Restricting the amount of characters per line in the translation tool may lead to translators abandoning subtitling altogether but adding guidelines that list good subtitling practices could solve some of the issues.

YouTube and its content evolve constantly as new people enter the platform. There is no telling which types of videos become popular within a year or two, and what type of people rise to the spotlight. One possible subject for a future study would be to examine the language used by vloggers and how they use it to engage with their audiences, for example by asking directly to write their opinions in the videos' comment section. The comments themselves could also prove an interesting study subject; one could, for example, study the language used in the comments, the discourse, interaction between viewers and content creators, to name a few.

Some translators do not want their name to show in the video description where the translator could be mentioned. It would be interesting, albeit difficult to study what is the reason they want to stay anonymous. It could be possible to try and collaborate with a YouTuber to get their attention and research their motivations. The phenomenon of invisible translators, and specifically the translators' will to stay anonymous are aspects of participant-oriented research that could result in an interesting case study, either in the YouTube environment or for example Netflix translators, as it is also a platform where the translators are very rarely given credit. Since fan translators are, at least to most extent, amateurs and lack the theoretical knowledge of professional translators, it would be interesting to study them at work; for example, what methods they use while translating, whether some of them resemble the methods of trained professionals and how they solve translation problems and issues.

WORKS CITED

Primary material

Dave Cad's YouTube channel (2018). Available at: https://www.youtube.com/user/Dave
        CadBlogs

What do we think about Finnish people? [YouTube video]. [Cited 16.2.2018]. Available
        at: https://www.youtube.com/watch?v=hsfOfWE5uX4

Subtitles from the following videos

AMERICA FIRST FINLAND SECOND REACTION [FIN/ENG Subs]
        Available at: https://www.youtube.com/watch?v=Q3UBGFdpt9U&t=22s

Am I A Feminist? (Snapchat Q&A!) | Dave Cad
        Available at: https://www.youtube.com/watch?v=kbX7ApOHs00&t=12s

British Guy VS LONGEST FINNISH WORDS! | Dave Cad with Cat Peterson
        Available at: https://www.youtube.com/watch?v=ybqJbk0dQ0U&t=25s

DRONE PROBLEMS
        Available at: https://www.youtube.com/watch?v=MjArAbXGy-8&t=7s

FINLAND & ENGLAND (The Differences) | Dave Cad
        Available at: https://www.youtube.com/watch?v=rdvTisZ6JQw&t=1s

LEARNING FINNISH | PART 1
        Available at: https://www.youtube.com/watch?v=eQYQGZhA2tw&t=
        16s

My Finnish Summer Begins! | Dave Cad
        Available at: https://www.youtube.com/watch?v=b6a-pN4denc&t=58s

MY VISIT TO TANZANIA (Serengeti & Zanzibar) | Dave Cad
        Available at: https://www.youtube.com/watch?v=xb3qB9pqVv4&t=18s

REACTING TO FINNISH MUSIC VIDEOS | Part 3
        Available at: https://www.youtube.com/watch?v=IHIqz9DUVOI&t=15s

REACTING TO FINNISH MUSIC VIDEOS | Part 4
        Available at: https://www.youtube.com/watch?v=6zr9yaeeDaM&t=16s

REACTING TO FINNISH MUSIC VIDEOS | Part 7
        Available at: https://www.youtube.com/watch?v=_U7SDfFA_n0&t=12s

REACTING TO FINNISH MUSIC VIDEOS | Part 8
Available at: https://www.youtube.com/watch?v=L-_-Xirihos&t=17s

Road Trip In Finland! | Part 1 | Dave Cad
Available at: https://www.youtube.com/watch?v=jvwkdtjHlK0&t=17s

TRYING TO SAY FINNISH TONGUE TWISTERS
Available at: https://www.youtube.com/watch?v=2tMs4BWxQ_k&t=16s

VIEWERS TEACH ME FINNISH
Available at: https://www.youtube.com/watch?v=r-XuEOyTElg&t=47s

100,000 SUBSCRIBERS Q&A
Available at: https://www.youtube.com/watch?v=8PAeEawJAvc&t=27s

2017 [FIN/ENG Subs]
https://www.youtube.com/watch?v=7AOMa9QqD2A&t=12s

Secondary material

Anastasiou, Dimitra & Gupta Rajat (2011). Comparison of crowdsourcing translation with Machine Translation. *Journal of Information Science* [Web document]. Issue 37(6) [Cited 23.10.2017], 637–359. Available at: http://journals.sagepub.com/doi/abs/10.1177/0165551511418760

Androutsopoulos, Janis (2013). Participatory culture and metalinguistic discourse: performing and negotiating German dialects on YouTube. In: Deborah Tannen & Anna Marie Trester (eds.). *Discourse 2.0: language and new media.* Washington, DC: Georgetown University Press. 47–71.

Boguki, Łukasz (2009). Amateur Subtitling on the Internet. In: Jorge Díaz Cintas & Gunilla Anderman (eds.). *Audiovisual Translation. Language Transfer on Screen*. Chippenham/Eastbourne: Palgrave Macmillan. 49–57.

Burgess, Jean & Joshua Green (2009). *YouTube. Online Video and Participatory Culture.* Cambridge Polity.

Business Insider (2015). *YouTube still doesn't make Google any money.* [Cited 3.10.2018]. Available at: https://www.businessinsider.com/youtube-still-doesnt-make-google-any-money-2015-2?r=US&IR=T&IR=T

Díaz Cintas, Jorge & Pablo Muños Sánchez (2006). Fansubs: Audiovisual Translation in an Amateur Environment. *The Journal of Specialised Translation* [Web document]. Issue 06 [Cited 2.10.2017], 37–52. Available at: http://www.jostrans.org/issue06/art_diaz_munoz.php

Díaz Cintas, Jorge & Aline Remael (2007). *Audiovisual Translation: Subtitling.* Manchester: St. Jerome Publishing.

Díaz Cintas (2008). Audiovisual translation comes of age. In: Delia Chiaro, Christine Heiss & Chiara Bucaria (eds.). *Between Text and Image. Updating research in screen translation.* Amsterdam: John Benjamins Publishing Company. 1–9.

Gambier, Yves (2008). Recent developments and challenges in audiovisual translation research. In: Delia Chiaro, Christine Heiss & Chiara Bucaria (eds.). *Between Text and Image. Updating research in screen translation.* Amsterdam: John Benjamins Publishing Company. 11–33.

Hauben, Michael (1997). Chapter 1 The Net and Netizens: The Impact the Net has on People's Lives [Web document]. In: Michael Hauben & Ronda Hauben (eds.). *Netizens: On the History and Impact of Usenet and the Internet.* Wiley-IEEE Computer Society Press. [Cited 8.1.2018], 3–34. Available at: http://www.columbia.edu/~rh120/ch106.x01

Hearts & Science (2017). *Who are the unreachables?* [Cited 8.1.2018]. Available at: https://www.hearts-science.com/forcesofchange/unreachables/infographic/

Helin, Irmeli (2008). Av-median ja kääntämisen kulttuuriset kontekstit [The cultural contexts of AV media and translation]. In: Irmeli Helin and Hilkka Yli-Jokipii (eds.). *Kohteena käännös. Uusia näkökulmia kääntämisen ja tulkkauksen tutkimiseen ja opiskelemiseen* [Translation as the target. New points of view to the study of and studying translation and interpretation]. Publications of the Department of Translation Studies 4. Helsinki: University of Helsinki. 131–148.

Immonen, Leena (2005). Sanat kuvatilassa [Words in the picture mode]. In: Hilkka Yli-Jokipii (ed.). *Kielen matkassa multimediaan. Näkökulmia kääntämisen tutkimiseen ja opiskelemiseen* [With language to multimedia. Points of view to the study of and studying translation]. Publications of the Department of Translation Studies 1. Helsinki: University of Helsinki. 165–177.

Immonen, Leena (2008). *Sanojen matkassa kuvatilaan* [With words to the picture mode]. Publications of the Department of Translation Studies 3. Helsinki: University of Helsinki.

Investor's Business Daily (2016). *This Is What Google Won't Tell You About YouTube's Financials.* [Cited 3.10.2018]. Available at: https://www.investors.com/news/technology/youtube-valuation-soaring-profits-blurry-as-facebook-amazon-loom/

Jenkins, Henry (P.I.), Ravi Purushotma, Alice J. Robison, Margaret Weigel & Katie Clinton (2007). *Confronting the challenges of participatory culture: Media education for the 21st century (Part One).* Nordic Journal of Digital Literacy. [Web document]. Issue 01/2007 Vol. 2 [Cited 20.2.2018], 23–33. Available at: https://www.idunn.no/dk/2007/01/confronting_the_challenges_of_participatory_

culture_media_education_for_the

Kieliasiantuntijat ry et al. (2020). *Käännöstekstitysten laatusuositukset 2020.* [Finnish subtitling quality recommendations 2020]. [Cited 11.2.2020]. Available at: https://kieliasiantuntijat.fi/wp/wp-content/uploads/2020/01/K%C3%A4%C3% A4nn%C3%B6stekstitysten-laatusuositukset-2020.pdf

Kotus (2018). *Kielitoimiston ohjepankki.* Available at: http://www.kielitoimistonohjepan kki.fi/

Kotus (2020a). *Ajanilmaukset: kellonajat (merkintä ja taivutus).* Available at: http://www.kielitoimistonohjepankki.fi/haku/kellonaika/ohje/143

Kotus (2020b). *Asteriski eli tähtimerkki (*).* Available at: http://www.kielitoimistonohje pankki.fi/haku/asteriski/ohje/12

Kotus (2020c). *Kohtelias kehotus tai pyyntö: konditionaali (Ottaisin kahvin, Veisitkö roskat).* Available at: http://www.kielitoimistonohjepankki.fi/haku/konditionaali/ ohje/500

Kotus (2020d). *Kolme pistettä.* Available at: http://www.kielitoimistonohjepankki.fi/hak u/ellipsi/ohje/47

Kotus (2020e). *Omistusliitteet: persoonapronimini ja omistusliite (laulaja ja hänen puolisonsa).* Available at: http://www.kielitoimistonohjepankki.fi/ohje/590

Kotus (2020f). *Passiivimaisia rakenteita: aina voit yrittää, "sinä-passiivi".* Available at: http://www.kielitoimistonohjepankki.fi/ohje/354

Kotus (2020g). *Pilkku*. Available at: http://www.kielitoimistonohjepankki.fi/haku/pilk ku/ohje/86

Kotus (2020h). *Predikatiivi: on mukava vai mukavaa, että tulet?* Available at: http://www.kielitoimistonohjepankki.fi/selaus/1488/ohje/601

Kotus (2020i). *Pronominit: joka vai mikä?* Available at: http://www.kielitoimistonohje pankki.fi/ohje/512

Kotus (2020j). *Verbi yksikössä tai monikossa.* Available at: http://www.kielitoimistono hjepankki.fi/selaus/2117/ohje/313

Kotus (2020k). *Yhdysmerkki yhdyssanassa, jossa erisnimi: Piritta-kahvila*. Available at: http://www.kielitoimistonohjepankki.fi/haku/niminen/ohje/100

Linna, Matti (2013). *Quality of Machine Translations by Google Translate, Microsoft Bing Translator and iTranslate4.* Unpublished Master's Thesis. University of Vaasa. English Studies.

Pedersen, Jan (2017). The FAR model: assessing quality in interlingual subtitling. *The Journal of Specialised Translation* [Web document]. Issue 28 [Cited 11.2.2020], 210–229. Available at: https://www.jostrans.org/issue28/art_pedersen.pdf

Pérez-González, Luis & Şebnem Susam-Saraeva (2012). Non-professionals Translating and Interpreting. *The Translator* [Web document]. Volume 18, issue 2 [Cited 2.10.2017], 149–165. Available at: http://www.tandfonline.com/doi/abs/10.1080/13556509.2012.10799506

Pérez-González, Luis (2014). *Audiovisual Translation: Theories, Methods and Issues.* Oxon: Routledge.

Pänkäläinen, Kristiina (2014). *This is daebak, here! – Fanitekstityksen piirteet* SHINee's One Fine Day *-ohjelmassa.* [*Characteristics of Fan Subtitles in* SHINee's One Fine Day]. Unpublished Master's Thesis. University of Tampere. Translation studies (English).

Saldanha, Gabriela & Sharon O'Brien (2014). *Research Methodologies in Translation Studies* [Web document]. Manchester: Routledge. [Cited 22.3.2018]. Available at: http://web.a.ebscohost.com/ehost/detail/detail?vid=0&sid=4d422194-9e4f-4c3a-a7ab-8615a864e25c%40sessionmgr4007&bdata=JnNpdGU9ZWhvc3QtbGl2ZQ%3d%3d#AN=756064&db=nlebk

Telegraph (2018). *The five best translation apps for travellers.* [Cited 25.4.2018]. Available at: https://www.telegraph.co.uk/travel/advice/The-five-best-translation-apps-for-travellers/

Tolson, Andrew (2010). A new authenticity? Communicative practices on YouTube. *Critical Discourse Studies* [Web document]. Volume 7, issue 4 [Cited 9.10.2017], 277–289. Available at: http://www.tandfonline.com/doi/abs/10.1080/17405904.2010.511834

Vehmas-Lehto, Inkeri (2005). Rakenneansoja ja piilovirheitä. Kohti selkeämpää käsitystä käännösvirheistä [Structural traps and hidden mistakes. Towards a clearer understanding on translation errors]. In: Hilkka Yli-Jokipii (ed.). *Kielen matkassa multimediaan. Näkökulmia kääntämisen tutkimiseen ja opiskelemiseen* [With language to multimedia. Points of view to the study of and studying translation]. Publications of the Department of Translation Studies 1. Helsinki: University of Helsinki. 49–74.

YouTube (2018). *About YouTube.* [Website]. [Cited 20.2.2018]. Available at: https://www.youtube.com/intl/en-GB/yt/about/

YouTube (2018). *Community Guidelines.* [Website]. [Cited 17.1.2018]. Available at: https://www.youtube.com/intl/en-GB/yt/about/policies/#community-guidelines

YouTube Help (2018). *Use automatic captioning.* [Website]. [Cited 10.10.2018]. Available at: https://support.google.com/youtube/answer/6373554

YouTube Red (2018). [Cited 12.1.2018]. Available at: https://www.youtube.com/red

YouTube statistics (2017). [Cited: 8.1.2018]. Available at: https://www.youtube.com/yt/about/press/

Appendix 1. Questionnaire form for the translators in Google Forms (Finnish)

Lyhyt kysely YouTube-kääntäjille

Tämä on kysely YouTube-käyttäjille, jotka ovat kääntäneet brittiläisen Dave Cadin videoita suomeksi ja suostuneet nimensä julkaisemiseen tekstitysten yhteydessä.

Kysely koostuu kolmesta osiosta: taustatiedoista, kieliosaamisesta sekä itse videoiden kääntämisestä. Kysymyksiin vastaaminen kestää n. 8 minuuttia, ja vastaukset käsitellään anonyymisti. Tuloksia tullaan käyttämään Vaasan yliopiston pro gradu -tutkielmassa. Voit vastata kaikkiin tai vain osaan kysymyksistä.

Taustatiedot

Koulutus-kohdassa valitse nykyinen opiskeluasteesi mikäli et ole vielä valmistunut. Ammatti-kohdassa kirjoita "opiskelija" mikäli et ole vielä valmistunut.

Ikä

○ <18

○ 18-25

○ >25

Koulutus

○ Peruskoulu

○ Lukio

○ Ammattikoulu

○ Yliopisto

○ Ammattikorkeakoulu

○ Muu: _____

Ammatti

Oma vastauksesi

Kielitaito

Muu kielitaito -kohdassa selvennä kielen perään taitotaso (esim. alkeet, kohtalainen, hyvä, erinomainen). Yksittäisten sanojen tai lauseiden osaamista ei lasketa kielitaidoksi.

Äidinkieli

○ suomi

○ ruotsi

○ Muu: _____

Vanhempien äidinkieli (jos eri kuin edellisessä)

Oma vastauksesi

Muu kielitaito

Oma vastauksesi

Oletko opiskellut kääntämistä?

○ Kyllä

○ En

Onko sinulla muuta kokemusta kääntämisestä?

Oma vastauksesi

YouTube-videoiden kääntäminen

Seuraavat kysymykset liittyvät YouTube-videoiden kääntämiseen, sekä Dave Cadin videoihin että niiden kääntämiseen yleensä.

Miten päädyit kääntämään Dave Cadin videoita?

☐ Huvin vuoksi

☐ Halusin kokeilla millaista kääntäminen on

☐ Halusin käytännön kokemusta kääntämisestä

☐ Muu:

Oletko tilannut Dave Cadin kanavan?

○ Kyllä

○ En

Millä perusteella valitset käännettävät videot?

☐ Sisältö

☐ Helppous

☐ Haastavuus

☐ Muu:

Millaisia ongelmia olet kohdannut kääntäessäsi videoita?

Oma vastauksesi

Jos vastasit edelliseen kysymykseen; miten ratkaisit ongelmat?

Oma vastauksesi

Oletko kääntänyt muita YouTube-videoita?

○ Kyllä

○ En

Jos vastasit edelliseen kysymykseen kyllä; millaisia videoita olet kääntänyt ja mille kielille?

Oma vastauksesi

Onko sinulla muuta yleistä kommentoitavaa kääntämisestä tai lisättävää tähän kyselyyn?

Oma vastauksesi

Appendix 2. English translation of the translator questionnaire

A short questionnaire for YouTube translators

This is a questionnaire for YouTube users who have translated the British YouTuber Dave Cad's videos into Finnish and agreed to have their username visible with the subtitles. The questionnaire consists of three sections: background information, language proficiency and the translation of the videos. Answering takes approximately 8 minutes, and the answers will be analysed anonymously. The results will be used in a Master's Thesis at the University of Vaasa. You may answer in all or part of the questions.

Background

In Education, pick your current status if you have not graduated yet. In occupation, write "student" if you have not graduated yet.

-Age: <18, 18-24, 25>
-Education: comprehensive school, upper secondary school, vocational school, university, university of applied sciences, Other?
-Occupation

Language proficiency

In Other language skills, define your language proficiency (e.g. basic skills, decent skills, good skills, excellent skills). Knowing single words or phrases does not count as language skills.

-Native language: Finnish, Swedish, Other
-Parents' native language (if different from previous answer)
-Other language skills
-Have you studied translation?
-Do you have some translation experience?

Translating YouTube videos

The following questions relate to Dave Cad's YouTube videos and to translating altogether.

-How did you end up translating Dave's videos? For fun, I wanted to try translating, I wanted translation experience, Other

-Are you subscribed to Dave Cad's channel? Yes, No

-How do you choose the videos you translate? Content, Easy to translate, Challenging to translate, Other

-What kind of problems have you encountered whilst translating?

-If you answered to the previous question; How did you solve those problems?

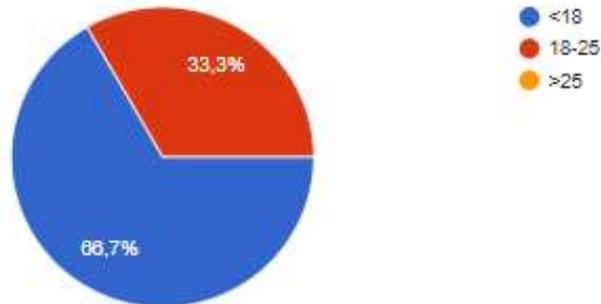-Have you translated other YouTube videos? Yes, No

-If you answered yes to the previous question; what kinds of videos have you translated and with which languages?

-Do you have anything else to comment on translation or add to this questionnaire?

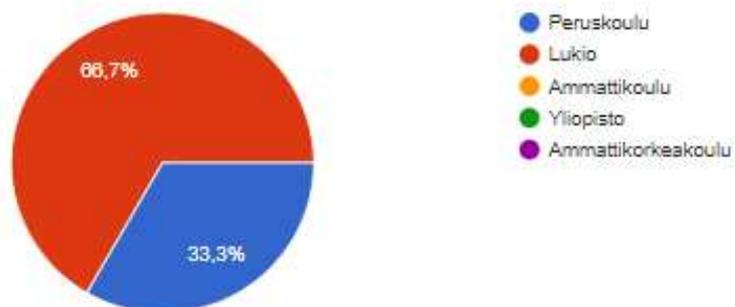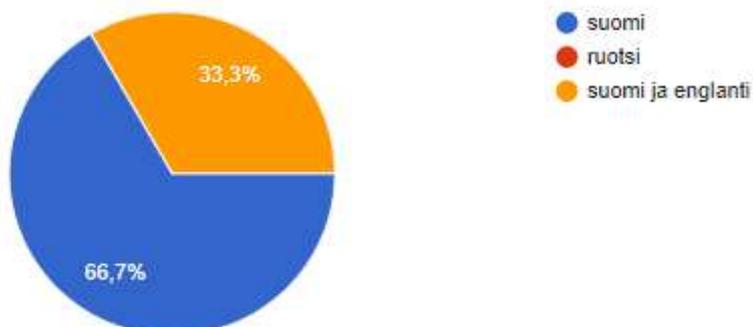Appendix 3. Questionnaire answers (Finnish)

Ikä

3 vastausta

- <18
- 18-25
- >25

33,3%

66,7%

Koulutus

3 vastausta

- Peruskoulu
- Lukio
- Ammattikoulu
- Yliopisto
- Ammattikorkeakoulu

66,7%

33,3%

Ammatti

2 vastausta

Kuvaaja

opiskelija

Äidinkieli

3 vastausta

- suomi
- ruotsi
- suomi ja englanti

33,3%

66,7%

Vanhempien äidinkieli (jos eri kuin edellisessä)

0 vastausta

Tähän kysymykseen ei ole vielä vastauksia.

Muu kielitaito

3 vastausta

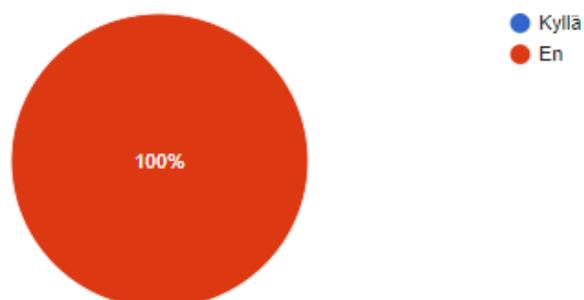Ruotsin kieliopin alkeet

englanti, ruotsi, saksa

englanti

Oletko opiskellut kääntämistä?

3 vastausta

- ● Kyllä
- ● En

100%

Onko sinulla muuta kokemusta kääntämisestä?
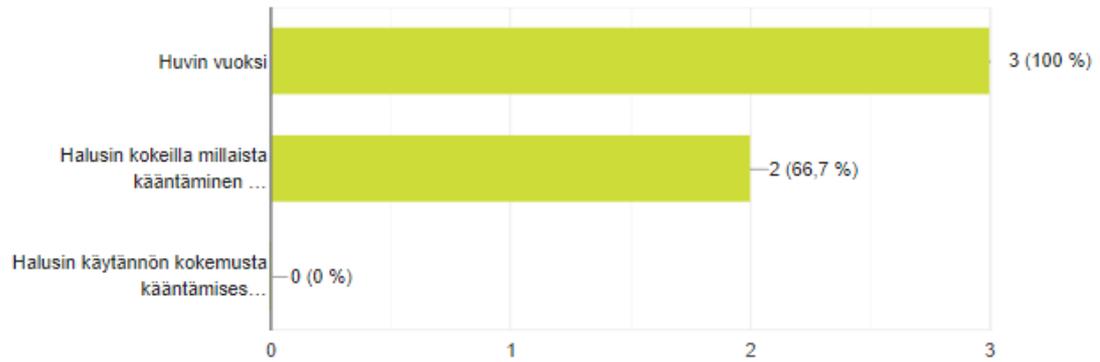
3 vastausta

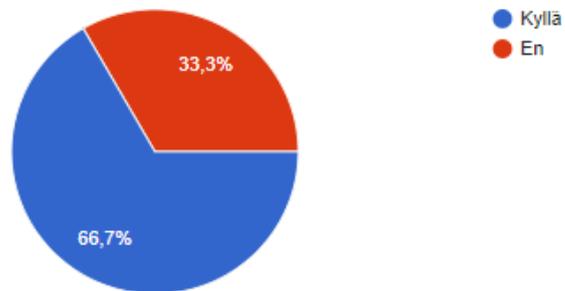| |
|---|
| Olen kääntänyt Discord nimisen alustan kääntämistä, sekä yhden botin joka toimii kyseisellä alustalla kääntämistä, sekä joskus loitin Scratchin kääntämistä mutta into silloin loppui lyhyeen. Omien videoiden kääntäminen |
| ei |
| Ei |

Miten päädyit kääntämään Dave Cadin videoita?

3 vastausta
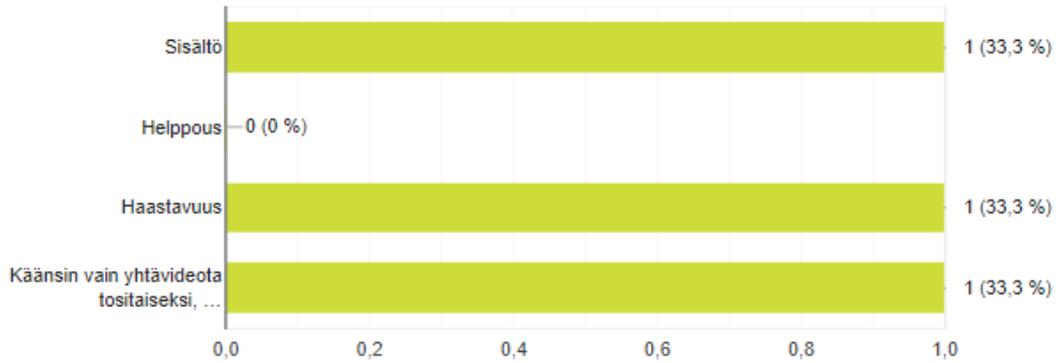

Huvin vuoksi	3 (100 %)
Halusin kokeilla millaista kääntäminen …	2 (66,7 %)
Halusin käytännön kokemusta kääntämises…	0 (0 %)

Oletko tilannut Dave Cadin kanavan?

3 vastausta

- Kyllä
- En

33,3%

66,7%

Millä perusteella valitset käännettävät videot?

3 vastausta

Millaisia ongelmia olet kohdannut kääntäessäsi videoita?

2 vastausta

Ajoittaminen(leikkauksen jump cuts), idiomaattiset sanonnat
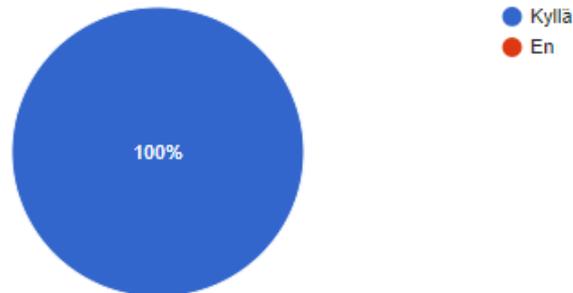
Empä oikeestaa mitää

Jos vastasit edelliseen kysymykseen; miten ratkaisit ongelmat?

1 vastaus

Ajoitukseen ei ratkaisua. Idiomaattisiin sanonoihin voi vain etsiä tai keksiä ratkaisua

Oletko kääntänyt muita YouTube-videoita?

3 vastausta

- Kyllä
- En

100%

Jos vastasit edelliseen kysymykseen kyllä; millaisia videoita olet kääntänyt ja mille kielille?

3 vastausta

Omia videoita suomi-englanti-suomi(Vancouver reissun videosta paloja suomennoksia). Suonennos tekstityksen tein myös "for disabled" eli esim kuurot voivat katsoa

samankaltaisia kuin Dave Cadin videot ja englannista suomeksi

Suomeksi ulkomaan kielisiä videoita

Onko sinulla muuta yleistä kommentoitavaa kääntämisestä tai lisättävää tähän kyselyyn?

1 vastaus

Hyvä kattava kysely.

Appendix 4. Questionnaire answers (English)

Age (3 answers): 66,7 % <18; 33,3 % 18-25

Education (3 answers): 33,3 % Comprehensive school; 66,7 % Upper secondary school

Occupation (2 answers): filmmaker/photographer, student

Native language (3 answers): 66,7 % Finnish; 33,3 % Finnish and English

Parents' native language (no answers)

Other language skills (3 answers): basic Swedish; English, Swedish, German; English

Have you studied translation? (3 answers): 100 % No

Do you have some translation experience? (3 answers): I've translated something for a platform called Discord, a bot that works on said platform, and once I tried translating something for Scratch but I didn't have enough interest to finish. Translating my own videos; No; No

How did you end up translating Dave's videos? (3 answers): 100 % For fun; 66,7 % I wanted to try translating

Are you subscribed to Dave Cad's channel? (3 answers): 66,7 % Yes; 33,3 % No

How do you choose the videos you translate? (3 answers): 33,3 % Content; 33,3 % Challenging to translate; 33,3 % Other: For now, I've only translated one video, but the intention was to translate others, as well

What kind of problems have you encountered whilst translating? (2 answers): Timing, jump cuts in editing, idiomatic expressions; Not really

If you answered to the previous question; How did you solve those problems? (1 answer): No solution for timing issues. You can only search or try to come up with a solution for idiomatic expressions

Have you translated other YouTube videos? (3 answers): 100 % Yes

If you answered yes to the previous question; what kinds of videos have you translated and with which languages? (3 answers): Similar to Dave Cad's videos, from English to Finnish; My own videos from Finnish to English and from English to Finnish (partial translations into Finnish in a video about a trip to Vancouver. I also made Finnish subtitles "for disabled" so that e.g. deaf people can view it; Translated foreign videos into Finnish

Do you have anything else to comment on translation or add to this questionnaire? (1 answer): Good, extensive questionnaire

Appendix 5. Errors in Finnish subtitles

Grammatical errors	Amount
unconjugated verb	13
unconjugated noun	38
unconjugated adjective/adverbial	4
wrong conjugation (verb)	29
wrong conjugation (noun)	29
wrong conjugation (adjective/adverbial)	8
wrong pronoun/errors with possessive pronouns	26
wrong conjunction	20
colloquial expressions	12
lacking a sentence element	15
mixing singular and plural expressions	5
Total	199

Linguistic errors	Amount
idiom translated falsely	10
word(s) or expression translated falsely	166
'it' translated directly	5
word-for-word incorrect translation	7
word/sentence not translated	41
interjection not translated	17
passive 'you' translated as active	2
translation uses different person	4
Total	252

Punctuation errors	Total
missing capital letters at the beginning of a sentence	218
missing capital letters in proper nouns	42
capital letter in the middle of a sentence	75
missing commas between main clauses	159
missing commas before a relative clause	96
missing commas before a subordinate clause	79
extra punctuation marks	29
missing dots	750
missing exclamation marks	8
missing question marks	11
other missing punctuation marks	106
incomplete ellipses	40
ellipses with too many dots	15
compound word written incorrectly	29
two separate words written as one	45
typing/spelling errors	64
numeral errors	14
missing a space between punctuation marks and following words	16
an extra space between punctuation marks and words	60
missing quotation marks	6
using an asterisk	8
Total	1870

Technical errors	Amount
too many characters per line	980
subtitles start too early	6
subtitles start too late	10
subtitles stay on-screen for too long after speech has ended	10
subtitles stay on-screen for more than 6 seconds	31
erroneous line break	31
subtitles too short	29
Total	1097