Subtitling the Internet
An investigation into subtitles for dynamic media

Sofia Dankis
Subtitling the Internet

An investigation into subtitles for dynamic media

Sofia Dankis

Abstract

This thesis will examine interlingual subtitles for online media. The research takes place within the descriptive translation studies paradigm. The source material consists of subtitles for a selection of webcasts found on the video sharing platform YouTube. The analysis will examine how the original dialogue has been translated as well as the form that the subtitles are presented in. An analysis of norm governed subtitles for television programs that follow established subtitling norms will also be made for comparison. Online media and broadcast media are divided into a dynamic and static category respectively in order to treat webcasts as an individual form of media. Dynamic media is defined as digital media created for the Internet, and static media is defined as media created for televised broadcasts. The results do not include qualitative assessments but have shown that subtitles for webcasts do not follow subtitling norms and display a wide range of unconventional approaches. The determining factor that illustrates the nature of webcast subtitles is attributed the diverse group of people writing them.

Keywords

Subtitling, translation theory, Swedish-English translation, audiovisual translation, online media, YouTube
Undertexter på Internet

En undersökning av undertexter för dynamisk media

Sofia Dankis

Sammanfattning


Nyckelord

Undertexter, översättningsvetenskap, svenska till engelska, medieöversättning, digital media, YouTube
Contents

1. Introduction ........................................................................................................... 1

2. Theoretical background ......................................................................................... 3
   2.1 Equivalence Theory ......................................................................................... 3
   2.2 Skopos Theory ............................................................................................... 4
   2.3 Localization ..................................................................................................... 4
   2.4 Audiovisual translation .................................................................................... 5

3. Webcast review ....................................................................................................... 9
   3.1 Webcasts and YouTube .................................................................................. 9
   3.2 Written text and YouTube ............................................................................. 10
   3.3 Subtitling and YouTube .................................................................................. 12
   3.4 Material .......................................................................................................... 13
      3.4.1 Group 1: Professional televised broadcast/professional subtitling .......... 14
      3.4.2 Group 2: Professional webcast/professional subtitling ....................... 15
      3.4.3 Group 3: Independent professional webcast/amateur subtitling .......... 15
      3.4.4 Group 4: Amateur webcast/amateur subtitling .................................. 16

4. Method ................................................................................................................... 18

5. Analysis .................................................................................................................. 20
   5.1 Quantitative data ............................................................................................ 20
   5.2 Group analysis ................................................................................................. 23
      5.2.1 Group 1: Professional televised broadcast/ professional subtitling ....... 23
      5.2.2 Group 2: Professional webcast/ professional subtitling ....................... 25
      5.2.3 Group 3: Independent professional webcast/ amateur subtitling .......... 28
      5.2.4 Group 4: Amateur webcast/ amateur subtitling .................................. 31
   5.3 General discussion ......................................................................................... 38

6. Conclusion .............................................................................................................. 41

Primary Sources ...................................................................................................... 43

Secondary Sources .................................................................................................. 44

Bibliography ............................................................................................................. 44

Filmography ............................................................................................................ 45
1. Introduction

There are many ways to approach the Internet when conducting an investigation. It contains endless information, has a multitude of contributors and there are infinite ways to use it. The vastness of the Internet makes it boundless, not only is there room for everyone to join the party but nothing ever really disappears either. When a text, image or video is uploaded online it remains there, to delete something completely is not as simple as burning every copy of a book or destroying the negatives from a roll of film. The world that exists online is forever moving and changing, it is dynamic. The world outside of the Internet is static in comparison. This difference is present in many different forms and areas but only one of these has been selected for this research. The specific focus is on subtitles, which is text that appears on video. So let us begin by looking at the different parts before we examine the whole.

Written text on the Internet differs from printed text in several ways. Besides reading off a screen instead of a piece of paper, the Internet does not require editing before publishing. Edited texts do exist online, both those that have previously appeared in print and original work. The keyword in this case is *require*. Texts that qualify as non-conventional in the sense that they do not adhere to the rules of grammar, punctuation, capitalization and so forth, run rampant on the World Wide Web. Not only is it possible for anyone with a connected computer to release a text online, they can take a variety of forms such as blogs, tweets, or comments. These texts can also theoretically reach every single person on the planet, unlike printed texts, which are only produced in limited editions.

Similar to text, online video is just as unruly. The videos that exist online vary from original content to pirated material, created by professionals and amateurs alike. There are videos that are only intended to be broadcast on the Internet, there are those that are meant for both worlds and then there are stolen videos that are shared illegally. Whatever their purpose, they enjoy the same freedom that written text does, they do not require an authoritative figure to allow for their release. In light of this, it is possible to differentiate online video from video that is broadcast on our televisions and screened in movie theatres in the same way that we distinguish between printed texts and digital texts, one side is static and the other dynamic.

The development of media has not only affected form but also milieu.

The television and the computer appear to be converging into one and the same screen offering very similar functions. They seem to be interchangeable. Today, we can watch the television on our PC or laptop and use the television set as a computer. Interactivity is a buzzword and its potential enormous (Díaz Cintas & Remael 2007, p. 142).

The merging of media platforms has created a new environment that has not yet been fully explored. Media that is made available online does not have to conform to the rigid standards of broadcast television; they belong to the dynamic world. This does not mean that all media on the Internet is in a
state of chaos, it means that its creation process now includes more choices. This dynamic media circumvents another boundary; information found online is not restricted by country but available on a global level. Some restrictions do exist, for example videos can be made unavailable for certain regions but the majority is open for everyone. This restriction is, however, a choice that is made by the party that uploads the video, it is not regulated automatically. Globally available information allows people from around the world to connect to each other and take part of new cultures. One of the tools that facilitate this connection is subtitling, breaking down the language boundaries whilst keeping the original art intact.

The combination of these two elements, text and video, brings us to the core of this thesis: subtitles. These strips of text that appear at the bottom of the screen have a number of functions and appear on all types of video. They can translate the dialogue into a different language, transcribe audio for a hearing impaired audience or act as a total supplement by adding information that is not present in the original at all. Subtitling is not a universal practice, nor is the balance between input and output languages even. “Onscreen, English is the all-dominant foreign language, and even major speech communities are turned minors in the process” (Gottlieb 2008, p.1). Subtitles in English are anomalous, Anglophone countries export far more video than they import. This means that their imports require a practice that Henrik Gottlieb calls “subtitling ‘upstream’, against the English current” (Gottlieb 2008, p.2). The subtitles that go against the current find themselves in a minority, much like subtitles for new media such as digital content only available online.

So how do subtitles act in this new world? Do they strive to remain as close as possible to the subtitles we see on television, the cinema and on DVDs? Or do they embrace the lawless country they find themselves in, replacing rules with anarchy? Subtitling for the Internet has already begun, out of the ten most viewed videos for 2013, eight of them included text and four were subtitled (YouTube Rewind 2013). The party has started, it’s time for the rest of the guests to arrive.

This research divides print media and digital media in two categories, static media and dynamic media. Printed text and televised broadcasts belong to the former and digital text and videos to the latter. The static category includes restrictions but benefits from training and structure. The dynamic is limitless but disorderly. These differences will be examined in the research by investigating a selection of subtitles from YouTube videos and televised broadcasts. All of the subtitles are translated against the current, from Swedish to English. The examination will look at strategies used in the translation of the original dialogue as well as the form and presentation of the subtitles. The aim of this research is determine which factors separate subtitles for online media from those for televised broadcasts.
2. **Theoretical background**

This section will identify the theoretical background for the data analysis. The study falls into the field of audiovisual translation, which is one branch in the translation studies field. An overview of relevant theories within translation studies will be provided, followed by a definition and close description of subtitling.

### 2.1 Equivalence Theory

Eugene Nida developed equivalence theory in 1964, an approach that examines whether the translator should preserve what is written or how it is written. Nida divided equivalence theory into two approaches, formal equivalence and dynamic equivalence. “Formal equivalence focuses attention on the message itself, in both form and content” (Nida 1964, p. 144). This approach preserves as much of what is written in the source text as possible, henceforth referred to as ST. Dynamic equivalence on the other hand states that “the relationship between receptor and message should be substantially the same as that which existed between the original receptors and the message” (ibid). This approach is more concerned with the way that the ST is written and attempts to emulate the tone rather than exact phrasing in the target text, henceforth referred to as TT.

Formal and dynamic equivalence both sacrifice one aspect of the ST in favor of another, neither approach produces an exact copy of the original. A TT of formal equivalence will resemble the ST as closely as possible, cultural references will not be replaced by those from the target culture, the structure will be as similar as possible and the grammatical rules in the target language will be the cause of any changes to word order. A TT of dynamic equivalence on the other hand will adapt the text for the target audience whilst keeping the core message, tone and character of the ST intact. This can involve replacing names and references, moving the text around in order to achieve a natural flow and making any changes necessary for the TT to feel the same way as the ST does. Each approach results in TTs that resemble the ST, but in different ways.

The difference between these approaches is significant to this research because they will be used as one of the key markers that highlight the difference between dynamic media and static media. Prior research has already concluded that televised broadcast tend to favor dynamic equivalence (Pedersen 2007, p. 7), this research will determine whether or not this is also true for webcasts.

This thesis refers to both dynamic equivalence and dynamic media, which are two separate concepts. For the purpose of this research, the term *dynamic equivalence* will reference Nida’s theory described in this section. The term *dynamic media* will refer to the classification of media outlined in chapter 1 and is not related to Nida or translation theory in any way.
2.2 Skopos Theory

Skopos theory examines the purpose and function of a translation. “The skopos of a translation is therefore the goal or purpose, defined by the commission and if necessary adjusted by the translator” (Vermeer 1989, p.200) Examining the skopos or purpose of a translation tells the translator how they should approach the source text, according to who the audience is, where it will be distributed and why is it being translated (ibid, p. 201). “Insofar as the duly specified skopos is defined from the translator’s point of view, the source text is a constituent of the commission, and as such the basis for all the hierarchically ordered relevant factors which ultimately determine the translatum” (ibid, p.192) A translatum is defined as “the resulting translated text” (ibid, p.191). Skopos theory is relevant to this research in that it allows us to incorporate another aspect of the translation process, namely the commission. “A great advantage of skopos theory is that it seeks to explicitate what the skopoi of translations actually are. In doing so, it emphasises the role of the translation commission. This can be seen as a statement about the skopos of a translation.” (Pedersen 2007, p.22) In several of the source texts the translator and commissioner are the same person, and this factor will be investigated in order to see how these texts diverge from texts with an external commissioner.

2.3 Localization

Localization is a process which fundamentally involves adapting a text or a product for a certain audience or market. It is often used for websites and video games, but can also be applied to television series and films. Localization is the second step of the globalization process, which allows one product to appeal to different groups of consumers on a cultural level. The first step in the process is internationalization, which involves neutralizing the product by stripping it of all cultural references. Localization entails adapting more than just the language; image and audio are also modified in order to fit into the target culture. Another difference between a translated text and a localized one is that a “localized text is not called on to represent any previous text; it is instead part of one and the same process of constant material distribution, which starts in one culture and may continue in many others” (Pym 2004, p.5).

Localization can be seen in certain areas of audiovisual translation, henceforth referred to as AVT, such as versioning which is the process which instead of “producing a film or TV programme and then translating it into various target languages (...) multiple-language versions of the same film or TV programme are produced” (Pedersen 2007, p. 30). This practice is often realized in countries that avoid other forms of AVT such as subtitling and dubbing, for example “the very small market shares of foreign films in Britain and the USA reflect that general Anglo-American audiences avoid non-Anglophone productions, no matter their genre or cinematographic qualities” (Gottlieb 2008, p. 1).
such countries, versioning is an attractive alternative. Recent examples of this include the Swedish films *Män som hatar kvinnor* (2009) and *Låt den rätte komma in* (2008) being remade into American versions, *The Girl with the Dragon Tattoo* (2011) and *Let Me In* (2010) respectively, only one or two years after their domestic release. The Swedish and Danish coproduction *Bron/Broen* (2011) has been remade twice, *The Tunnel* (2013) is set in England and France and *The Bridge* (2013) takes place in the US and Mexico.

There are several different techniques that can be employed during the localization process. Hybrid strategies combine different strategies instead of favoring one over the other, allowing for cultural customization (Pym 2010a, p.5). Pym claims that hybrid strategies “move some communicative decision-making to the user’s side, thus constituting one of the truly liberating and democratic aspects of electronic communication.” (ibid, p.5). This approach can include linking additional elements to the text, for example hyperlinks to other websites or embedded applications. These elements can be compared to footnotes in written text, only these can include other features such as image and audio.

Comparing websites to print media separates the two into a dynamic category and a static category. It should be noted that these categories are applicable to the final product, print media can be transformed countless times before it is published. Therefore when discussing printed text, it is the finished paper copy that is referenced to. Websites, however, can be updated rather than being replaced by a new version, making them unfixed and interactive texts. A website can be altered through coding as an alternative to a human translator producing a text from scratch. This allows the components that need localizing to be singled out and altered by using translation memories and glossaries.

The distinction between these two media is significant because this difference is also applicable to televised broadcasts and web broadcasts. Televised broadcasts have temporal and spatial restrictions, but also restrict viewers from interacting with the program or take advantage of any additional tools. This places televised broadcasts in the static category together with print media. A web broadcast, however, allows the viewer to access a wider range of options which places it in the dynamic category.

### 2.4 Audiovisual translation

The previous sections in this chapter have established that the shift from one language to another will alter the content in varying degrees depending on the approach used by the translator. This section will examine the impact this shift has on translated polysemiotic media, where the source text is made up of more than one code.

Audiovisual translation includes “all forms of translation that appears in an audio and/or visual context” (Pedersen 2007, p.30). The main forms of AVT are subtitling, dubbing and voice-over (Díaz
Cintas & Remael 2007, p.8). European countries usually only use one of the forms for televised broadcasts and cinematic releases (Pedersen 2011, p. 4). Dubbing replaces the original audio soundtrack with a translated one whereas voice-over keeps the original soundtrack and adds additional narration on top of it in a different language (ibid, p. 4). The term **subtitling** is defined as “a translation practice that consists of presenting a written text, generally on the lower part of the screen, that endeavors to recount the original dialogue of the speakers, as well as the discursive elements that appear in the image (…) and the information that is contained on the soundtrack” (Díaz Cintas & Remael 2007, p. 8). In order to remain consistent, the term **subtitle** will be used instead of **caption**, which the practice is also known as, and also the term that is used on YouTube.

Subtitling can be pre-prepared or live, open or closed, interlingual or intralingual. Pre-prepared subtitles are, like the name implies, produced before the program is aired. Live subtitles are used for live programming, such as news broadcasts and award shows (ibid, p.19). Open subtitles “are burned or projected onto the image and cannot be removed or turned off (…) [closed subtitles] can be added to the programme at the viewer’s will. The subtitles are hidden and can only be seen with an appropriate decoder or when the viewer activates them on DVD” (ibid, p. 21). Open subtitles are often found on VHS tapes and cinematic releases, closed subtitles are more common for DVD. Televised broadcasts make use of both open and closed subtitles.

Intralingual subtitling “involves a shift from oral to written but stays always within the same language” (ibid, p. 14). Intralingual subtitling can also be used as a language-learning tool. Other uses include clarifying what the speaker is saying, for example in cases where heavy dialects are involved, there are other auditory channels interfering or the speaker is mumbling or whispering. Karaoke also makes use of intralingual subtitling, allowing anyone watching to sing along without missing any of the lyrics (ibid, p.16-17).

Interlingual subtitles translate the source language to a different target language. “Interlingual subtitling is unique in that the message is not only transferred from one language to another, but also from one mode to another: from the spoken mode (usually) to the written mode” (Pedersen 2011, p.11) Using interlingual subtitles to translate the verbal dialogue into other languages bridges the gap between countries by allowing media to be shared between them. “Watching and listening to films and programmes subtitled from other languages helps us not only to develop and expand our linguistic skills, but also to contextualize the language and culture of other countries” (Díaz Cintas & Remael 2007, p. 15). Subtitling can also open up a door for the hearing impaired. This type of subtitling is known as SDH (Subtitling for the deaf and hard-of-hearing) (ibid, p. 14) In these cases they “also incorporate all paralinguistic information that contributes to the development of the plot or the creation of atmosphere” (ibid, p. 14). These subtitles can be activated through teletext in some European countries but can also be transmitted on a separate line in other areas (ibid, p. 14). SDH can be both interlingual and intralingual.
Whether or not subtitles are intralingual or interlingual, they will always translate one code into another.

Literary translation and interpreting, the two traditional counterparts in interlingual communication, are horizontal types, moving in a straight line from one human language to another, without shifting language mode: speech remains speech, and writing remains writing. Subtitling, on the other hand, can be either vertical or diagonal. Being intralingual, vertical subtitling limits itself to taking speech down in writing, whereas diagonal subtitling, being interlingual, ‘jaywalks’ (crosses over) from source-language (SL) speech to target language (TL) (Gottlieb 2001a, p. 17)

Any type of subtitling is seldom a verbatim version of the dialogue. This is caused by temporal and spatial constraints. One way to work around these obstacles is to take advantage of the visual and non-verbal acoustic channels. “Audiovisual programmes use two codes, image and sound, and whereas literature and poetry evoke, films represent and actualize a particular reality based on specific images that have been put together by a director.” (Díaz Cintas & Remael 2007, p.9). The audience takes in information from different sources by watching the action, listening to the audio and reading the subtitles. If one source is forced to omit information, another one can be used to balance the full equation. For example a scene depicting an argument can omit expletives or proclamations of anger. The audience can see and hear that the characters are angry which allows the translators to focus on information that is important to the plot. These channels can also affect the approach taken by the translator.

The continual presence of the other semiotic channels (the image and the ‘international’ music & effects track) in translation means that on the axis ranging from strictly verbatim rendition of the original (verbal) text to free, target-culture recreation of the text, e.g so-called localization, translations of commercial film and TV productions would be expected to stay near the source-oriented pole. (Gottlieb 2008, p. 3-4)

The temporal and spatial constraints causes subtitling to put the source text through the following processes: cueing, segmentation and condensation. These steps are taken to ensure that the written text is synchronized with the audio and image. Subtitling “is constrained by the respect it owes to synchrony in these new translational parameters of image and sound (subtitles should not contradict what the characters are doing on screen), and time (i.e. the delivery of the translated message should coincide with that of the original speech)” (Díaz Cintas & Remael 2007, p. 9). In order to adhere to synchrony, a subtitler follows subtitling norms and as well as in-house guidelines provided by the company commissioning the work (Pedersen 2011, p.22). Cueing determines “the precise moments when a subtitle should appear on screen – known as in-time – and when it should leave the screen – known as out-time – according to a set of space and time limitations” (ibid, p.30). The segmentation process divides the verbal dialogue in subtitles as well as splits the subtitles into one or two lines. This process separates subtitles from other forms of written texts. “Contrary to the way we consume a written text on a page, the compulsory segmentation of subtitles gives them a fragmentary nature that complicates their reading. Each subtitle is an isolated unit, disconnected physically from the preceding and the following subtitles.” (ibid, p. 103) Cueing and segmentation embody the temporal and spatial constraints of subtitling, respectively. A direct product of these constraints is condensation, “all the
verbal content in the ST cannot always be represented in the subtitles. Something usually has to be left out, or rather: the remaining TT message has to be edited so as to (ideally) say the same thing, in fewer words (or at least in fewer characters), as the ST” (Pedersen 2007, p. 45). Editing out parts of the dialogue in the subtitles is known as text reduction. “Partial reduction is achieved through condensation and a more concise rendering of the ST. Total reduction is achieved through deletion or omission of lexical items” (Díaz Cintas & Remael 2007, p. 146).

Subtitlers know when and where to apply these processes by following subtitling norms. The term subtitling norms will for the purpose of this research refer to established subtitling standards. Gideon Toury describes norms as “a favoured mode of behaviour within a heterogeneous group” and places them between rules and idiosyncrasies (Toury 1995, p.54). This means that norms are not as definite as rules, but are instead known within a certain group or field to be the preferred choice.

Subtitling norms include the following points:

- On-screen duration: “This limit, of some twelve subtitle characters per second (12 cps), means that a full two-liner should stay on the screen for at least five seconds” (Gottlieb 2001a, p.20).

- Character limits: “a full line of contemporary subtitling in the present study is considered to have 36 characters, and a full two-liner 72 characters” (Pedersen 2011, p.19).

- Punctuation: Dashes are used to indicate more than one speaker in the same subtitle (Díaz Cintas & Remael 2007, p. 111).

- Font: Italics are used to reference dialogue heard off-screen (ibid, p. 124). They are also used for songs, which are also written out so each new line begins with a capital letter, names of songs are usually reserved for SDH (ibid, p. 127).

Subtitling norms are imposed for televised broadcasts because of the professional, or static, atmosphere they appear in. The entertainment business is just that, a business. That demands that professionals create each aspect of a televised broadcast, which includes the subtitling. Professional subtitlers have been taught what these norms are and have been trained to follow them. Webcasts on the other hand do not appear in a strictly professional atmosphere, but are broadcast on a dynamic platform that is open for everyone. This means that anyone can be involved in any aspect of a production, whether they are trained for it or not. It is in this light that the data for this research will be examined.
3. Webcast review

This section will provide a context for the analyzed data itself by examining the environment that it appears in, which is the video-sharing website YouTube. The nature of webcasts, an overview of how text is used for this type of media and a description of YouTube’s own subtitle function will be provided. The source material will also be presented here.

A webcast is defined as “a transmission of sound and images (as of an event) via the World Wide Web” (Merriam-Webster 2014). For the purpose of this research the term webcast will refer to video that is only released online without the intention of being aired on network television. This term will in this sense only apply to video that is available globally, meaning that media available online but only to a restricted audience (for example paid subscribers or residents of a specific country) will not fall into this category. The term user will refer to the party responsible for uploading a video onto YouTube, the term viewer will refer to the audience of online media. The term televised broadcasts will be used to refer to video that is primarily viewed on network television, restricting the audience to those within the limits of a specific country.

3.1 Webcasts and YouTube

The video streaming site YouTube is described by Jean Burgess and Joshua Green as being:

variously understood as a distribution platform that can make the products of commercial media widely popular, challenging the promotional reach that the mass media is accustomed to monopolizing, while at the same time a platform for user-created content where challenges to commercial popular culture might emerge, be they user-created news services, or generic forms such as vlogging – which might in turn be appropriated and exploited by the traditional media industry.

(Burgess & Green 2009, p.6)

YouTube videos have been selected for this research in part to narrow the source material. YouTube was specifically chosen because it is an open and free source unlike video streaming sites that require paid subscriptions, such as Hulu. Burgess & Green liken Hulu to static media since it functions more like television than a website (Burgess & Green 2009, p. 32). They go on to state that it lacks “the civic opportunity of YouTube that responds to many of the ways media consumers use content in their everyday practice” (ibid, p. 32) Since the time that this argument was made there has been an increase in websites like Hulu, for example Netflix, HBO GO and Amazon Prime. These services also function more like static media than dynamic media despite their online availability. They require the viewer to have a membership, pay a fee and restrict content according to region which is why the source material does not include any webcasts available through these services. By selecting the alternative that is described as both civic and an everyday occurrence, the webcasts will only include dynamic media and have a broader scope which includes professionals and amateurs alike.
The distribution of webcasts is one of the main factors that set them apart from televised broadcasts. Televised broadcasts can only reach a finite number of audience members, for example a local newscast will only be available to the population of a certain area. Webcasts on the other hand can spread on a global level and can be viewed multiple times, when they reach a very large audience they are known as viral videos. Jean Burgess states “the term ‘viral video’ is used to refer simply to those videos which are viewed by a large number of people, generally as a result of knowledge about the video being spread rapidly through the Internet population via word-of-mouth.” (Burgess 2008, p.1) The word-of-mouth element offers some insight into the incentive behind professionally produced media being released on YouTube. This type of distribution is possible for dynamic media since the videos exist digitally and not just as physical copies. When a video is uploaded to YouTube it is given its own webpage with a specific web address. This allows viewers to share a direct link to the video through various platforms, such as social media websites, email and blogs. If one were to share static media, it would require a physical copy or instructions for the other party to view a specific channel at a specific time.

According to Burgess & Green (2009, p.2), one of the more popular myths that depict how YouTube gained popularity was through the viral spread of a sketch video from the television program Saturday Night Live (1975-?). The video was eventually taken down due to copyright infringement because it had been uploaded by a party that did not own it. This not only demonstrates one of the key usages of the website, an easy platform to watch and share popular culture, but also illustrates the evolution of the website as today’s networks are uploading clips from their shows themselves, including Saturday Night Live. This allows viewers to sample shows, and by sharing those clips they are marketing the show for the network.

YouTube is an example of how dynamic media allows for more contributors and a wider distribution than any static media platform can offer. By taking these factors into account, we can see that the webcasts examined in this research have been created and distributed under a different set of circumstances than the televised broadcasts. These circumstances have allowed for webcasts to include a wide and diverse range of creators, made up of both professionals and amateurs. It is therefore possible to assume that webcast subtitlers are also a more diverse group than their static media counterparts. This question will be taken into consideration in the analysis of the source material.

3.2 Written text and YouTube

Text is used in a variety of ways on YouTube, subtitling is only one of them. In order to place the subtitled media that will be examined in this research in it’s proper context, the following overview will outline different ways that text is used in YouTube videos.
Interlingual and intralingual text is used for subtitling, it is possible to find videos with either closed or open subtitles. There is a great deal of pirated material on YouTube which includes open subtitles from the original broadcast. These are not considered to be subtitles for webcasts because they were originally written for and broadcast on television. Whereas interlingual subtitles translate the dialogue into different languages, intralingual subtitles are often used for comedy as well as a clarification aid for low quality videos. An example of these types of subtitles are found in the videos uploaded by YouTube user pewdiepie, who has the most subscribed to channel of all time at this date. At the time of this writing the channel has over 23 million subscribers. Felix Kjellberg, pewdiepie’s given name, began uploading video recordings of himself playing video games. His constant commentary, screaming and creative editing have gained a following, his fans calling themselves “bros” (Gallagher 2013). Kjellberg adds unconventional subtitles to his videos, the words appearing as they are spoken in different open-faced, colored font. The subtitles are in English, except for in the cases when Kjellberg exclaims a few words or sentences in Swedish. They do not always relay exactly what is spoken, at some points they describe non-verbal audio and some instances comment on the action. The language utilizes a great deal of Internet jargon and slang, as well as symbols replacing words at some points.

Text can also be used as a learning tool. Tutorials and instructional videos are very popular on YouTube and text is often used to highlight steps and instructions. The viewer can see what they should be doing, but the text remains on the screen longer than a verbal utterance, allowing them to take their time and focus on the task at hand instead of trying to remember what they were just told to do.

Intralingual text is also used in what is called a lyric video, a new trend that merges music videos with karaoke videos. The only thing displayed on the screen are the lyrics to the song as they are being sung, much like a karaoke video. Lyric videos were made popular on YouTube, and according to Emily Zelmer it was Cee Lo Green’s song “Fuck You” that began the trend (Cee Lo Green 2010). The main appeal that this type of video holds is that the cost is considerably lower than a traditional music video. Instead of a production requiring numerous employees, from crewmembers to the artist themselves, lyric videos can be made by one person and a computer. Zelmer also points out that the “popularity of lyric videos also corresponds to sales, giving labels an early indicator of a hit song” (Zelmer 2013).

The interactive feature of YouTube videos utilizes text for annotations and advertisements. The annotations function allows users to add text and links to their videos. They are commonly used to encourage viewers to subscribe to channels, watch the next video or like the clip. It is possible to add banner advertisements to YouTube videos, which is one of the ways that a user can make money on their videos. The banners appear at the bottom of the video, some of them can be turned off by
switching off the annotations function but others need to be clicked on to shut down like a regular pop-up window.

Taking the different types of ways that written language is used on YouTube into consideration, it is possible that this phenomenon has created a willingness to read while watching. If webcasts continue to develop and become a larger part of the media landscape, it is possible that viewers will adapt to them in the same way that audiences of televised broadcasts have. Research on subtitling for static media has exhibited that audiences in areas where they are the main form of AVT have adapted to reading and watching simultaneously.

With the domination of the Anglophone entertainment industry, it means that a great deal of what you will encounter in cinemas and on television will be English-language films and TV programmes with subtitling in the local language. This means that citizens of these countries spend a not inconsequential part of their day reading subtitles. (Pedersen 2011, p. 1)

The viewers of dynamic media are not restricted by their location in the same way that static media audiences are, they are only limited by their personal interest and taste in entertainment. Therefore, if exposure is one of the factors that contribute to the acceptance of subtitles it is possible that the dynamic media viewer has already begun to grow accustomed.

3.3 Subtitling and YouTube

Subtitles can be added to any YouTube video, the official help page motivate their use by stating that they “open up your content to a larger audience, including deaf or hard of hearing viewers or those who speak other languages” (Google 2014). Another benefit that subtitling a YouTube video has is that the subtitle transcript is included when search engines generate results. “Closed captions are a positive signal to the search engines to help them better understand the video content – and return that content in the blended search results for a given search phrase, particularly the more long tail ones.” (Marshall 2012). A user can upload either a caption file or a transcript file, and in the latter case the viewer will also be able to access the entire transcript. When viewing a video, the option to include subtitles can be selected from the setting menu that also includes video quality and the option to turn advertisements on and off. If a video includes subtitles an icon that reads “CC”, which stands for “closed captions”, appears underneath the title.

YouTube has created a straightforward process for users to subtitle their videos, both by the simple interface but also by including the function to request a translation. The user is given the option to request a translation through a vendor resulting in a fee, or by sending out an invitation to personal contacts. The official help page also refers to external services that provide subtitling software for YouTube videos. Subtitles can also be added by using voice recognition technology, which often results in faulty translation. An example taken from a pewdiepie video is the translation of “I freakin’ love Fridays” into “African Fridays” (pewdiepie 2010).
YouTube does not highlight any of the options; they present the user with a choice and make every step along the way as simple as possible. The choice to add subtitles is set on the same level as selecting the video quality. Their subtitling service indicates that subtitles are becoming more available and accepted.

The subtitle function also illustrates one of the ways that participatory culture, “a term that is often used to talk about the apparent link between more accessible digital technologies, user-created content, and some kind of shift in the power relations between media industries and their consumers” (Burgess & Green 2009, p. 10), can improve and impact the spread of uploaded videos. A user cannot add a subtitle file to another user’s video, but they can submit them. A user can also very simply request this service, and the vastness of the Internet allows them to reach further than trained professionals who would charge a fee. The option to rely on volunteer work and fan contributions is open to both private citizens and professional companies.

At the time of writing, there are no established norms for webcast subtitles. The new format brings new technology along with it. The effect that the arrival of the DVD had on subtitling was described as “not only is professional practice changing but the very essence of subtitling and the conventions applied are also in flux” (Díaz Cintas & Remael 2007, p.139). This comment is in regard to the available technology that came with DVD, for example closed subtitles instead of open and new accessible fonts, colors and symbols. These new factors resulted in “new conventions that diverge acutely from what has been considered standard practice in interlingual subtitling up until now have started to crop up in subtitled programmes commercialized on DVD” (ibid). Considering the impact that DVD made on subtitling, it is conceivable that webcasts can have a similar effect. In this case new technology is available, for example utilizing hyperlinks in the same way as they are used during the localization process. YouTube’s own captioning service allows the viewer to customize color, font, position and size. Subtitles for webcasts have handed even more power over to the viewer. The DVD allowed the viewer to decide whether they wanted subtitles or not and provided a limited number of languages to choose from. The webcast has furthered this by giving the viewer even more power by allowing them to customize the subtitles and even get involved in the translation stage through crowdsourcing.

3.4 Material

This section presents the material for this research. The analyzed data consists of nine webcasts found on YouTube and two televised broadcasts that are taken from their DVD releases. The polysemiotic nature of the material allows for the analysis to be based on the action taking place on screen as well as two transcripts for each video, one of the verbal dialogue and the other of the subtitles. For the purpose of this research the transcribed dialogue will act as the ST and the subtitles as the TT. The on-
screen action will be taken into account in the analysis but will not be referred to as the ST since it technically does not qualify as text.

The data has been organized into four groups:

- Professional televised broadcast/ professional subtitling
- Professional webcast/ professional subtitling
- Independent professional webcast/ amateur subtitling
- Amateur webcast/ amateur subtitling

The data has been sorted into these groups in order to differentiate between the conditions that each video was created under. It is possible that these differences will affect the results and it is necessary to take them into account before a general evaluation is made. The analyzed webcasts are presented and also compared to possible televised broadcast counterparts. The range of webcasts covers professional production with professional subtitling, professional production with amateur subtitling and amateur production with amateur subtitling. Pirated material has not been included in the source material.

The term professional webcast refers to media produced by professionals on behalf of a corporation. The term independent professional webcast refers to media produced by professionals without corporate support. The term amateur webcast refers to media produced by non-professionals. The term amateur subtitling signifies a text produced by a translator without formal training within the discipline. The term professional subtitling signifies a text produced by a trained translator on behalf of a company or translation agency. As the research is not qualitative these terms are neutral and used without prejudice.

**3.4.1 Group 1: Professional televised broadcast/professional subtitling**

These broadcasts are taken from the DVD releases of two Swedish television series, which originally aired on Swedish television. They are categorized as static media. They will act as the control group and be used as a comparison in order to highlight the ways in which the web broadcasts differ from televised broadcasts.

**3.4.1.1 Sandhamn**

*Morden i Sandhamn* (henceforth referred to as *Sandhamn*) is a criminal drama that was originally aired on the Swedish channel TV4 in 2010. The subtitles were written by Helena Johansson for Hippeis Media. The series has been running for three seasons so far and is made up of three episodes per season. The action takes place in Sandhamn, an area in the Stockholm archipelago, and revolves around a murder mystery. The selection for this research was taken from the second episode of the first season. The seven-minute sample spans three consecutive scenes containing mainly dialogue.
3.4.1.2 HippHipp

HippHipp! Itzhaks Julevelgrium (henceforth referred to as HippHipp) is a special edition of the television series HippHipp! which originally aired on Sveriges Television. The original series is made up of sketches but this has the form of a traditional, scripted mini-series. This particular edition is a Christmas special which follows a variety of characters in different areas in Sweden as they prepare for the holidays. The four-minute selection spans two consecutive scenes with few breaks between the dialogue.

3.4.2 Group 2: Professional webcast/professional subtitling

This group is made up of webcasts that have been produced on behalf of professional production companies or corporations. The subtitles have been provided by professional translation agencies. They are categorized as dynamic media.

3.4.2.1 Jaharà

Alex & Sigge I Jaharà: Anställningsintervjun [with subs] (henceforth referred to as Jaharà)

This webcast is a sketch comedy series produced by Hard Hat Media in 2013. The subtitles were written by Elin Larsson for SDI Media. The short episodes are around 5 to 10 minutes and are unrelated to each other. Subtitles were not added to the first sketches, but were a later addition and include English and SDH.

This production can be likened to a comedy series with shorter sketches, such as Saturday Night Live (1975-?).

3.4.2.2 Rohdi

Rohdi Heintz, från from Guldknappen 1981-2011 på at Nordiska museet (henceforth referred to as Rohdi) is a recording of an interview with artist Rohdi Heintz in 2011. The video is produced by Nordiska Museet and the subtitles were provided by Comactiva translations. An interviewer is heard ask a question in the beginning of the video but the rest is made up of Heintz describing his work.

This production can be likened to a televised interview for an informative program or a news report.

3.4.3 Group 3: Independent professional webcast/amateur subtitling

This group is made up of webcasts that have been created by independent professionals which means that the creators have a certain level of training and skill but the webcast has not been commissioned or funded by an external party. The subtitles have been written by a person without formal training within translation or subtitling. They are categorized as dynamic media.

3.4.3.1 Sanning

Sanning eller Död (Kortfilm/Short film) (henceforth referred to as Sanning)
This webcast is a short film produced by Lineup Films in 2011. The film’s Swedish producer Tom Grejs wrote the subtitles for Sanning with the help of the American screenwriter (Grejs, T. 2014. personal communication, February 27). The story takes place during one evening on a cliff in Dalarna where two young people challenge each other to tell the truth or jump off a cliff.

This production can be likened to a drama series such as The O.C. (2003-2007).

3.4.3.2 BOB

BOB (Kortfilm / Short film) ENG-subs available! (henceforth referred to as BOB) is a short film produced by Hedefors Productions from 2013. It was submitted to a film festival but is otherwise only available online. The subtitles for this film were written by the directors, both of them non-professional translators. Their motivation behind adding subtitles was so that a friend of one of the actors would be able to understand it, they had no intention of attempting to reach a global audience (Edstam, M. 2014, personal communication, March 3). The film centers around a house cat who prefers to stay indoors, all of the animals are played by people.

This production can be likened to Dr. Seuss’ The Cat in the Hat (2003).

3.4.3.3 Sneakerliv

Sneakerliv is a documentary by Leif Egnell from 2013 about sneaker enthusiasts and collectors. The subtitles were written by Egnell himself, who described the subtitling process as first transcribing the dialogue and then using Google Translate to translate it into English (Egnell, L. 2014, personal communication March 21). The documentary is made up of interviews with eight different people who discuss their love for sneakers, why they started collecting and show off some of their shoes.

This production can be likened to a documentary series, for example Vice (2013-?).

3.4.4 Group 4: Amateur webcast/amateur subtitling

This group is made up of webcasts that have been created and subtitled by amateurs, meaning that the webcasts are a side project or hobby rather than a profession or source of income. They are categorized as dynamic media.

3.4.4.1 vlog

Jag pratar Svenska/ A Swedish vlog (henceforth referred to as vlog) is a video blog made by a woman named Alyona under the screen name The Sleep of Reason in 2014. Neither English nor Swedish are the mother tongue for the speaker/translator who is originally Russian. In her video she explains that she is trying to learn Swedish, so she created a video blog in order to practice but added the subtitles so that her followers would be able to understand her.

This production can be likened to a reality series such as The Real World (1992-?).
3.4.4.2 Gathering

Stockholm YouTube Gathering 2013 – Vad hände? (henceforth referred to as Gathering) is a montage of clips from the event Stockholm YouTube Gathering from 2013. It was created by Martin Arvebro and it is assumed that he provided the subtitling. The video features clips from panel discussions, interviews and performances as well as clips from a group outing to a park in Stockholm.

This production can be likened to a home video compilation from a family vacation or party.

3.4.4.3 Mealtime

Chop Chop Carnage Stew (Alternative Pyttipanna) – Regular Ordinary Swedish Meal Time (henceforth referred to as Mealtime) is a popular web broadcast, this particular video from 2011 had over two million views at the time of this research. The only information obtained regarding the creators of the video is taken from their Swedish Wikipedia page, which indicates that it was started by Niklas Odén, Anders Söderman och Tom Brännström (Wikipedia 2013). The video involves a man yelling a recipe for the Swedish dish “pyttipanna” whilst throwing ingredients and cooking violently. It should be noted that the dialogue is mainly delivered in Swedish but there are a few instances of words or full sentences in English. This mix of two languages is stated in the title by the addition of “(In Swenglish)” (Mealtime 0:22).

This production can be likened to a cooking show, for example The Naked Chef (1999-2001).

3.4.4.4 Haffa

Haffa Guzz 2.0 (henceforth referred to as Haffa) is a video of a young man walking around an area in central Stockholm called “Plattan”. The video was uploaded in 2013 by user RandomMakingMovies which appears to be a group of young men but I was unable to find further information about them. The video follows the young man as he approaches different young women and asks for their phone numbers, for them to kiss him or for them to take his virginity. It is apparent that all of the parties other than the young man are unaware that they are being filmed.

This production can be likened to prank shows, for example Punk’d (2003-2012).
4. Method

This section will present how the source material presented in the previous section will be treated. The different translation theories and strategies presented in Chapter 2 will be used to analyze content of the subtitles and the section regarding AVT and subtitling norms will analyze the subtitles form. Quantitative data will be extracted and presented in a table, this information will be considered regarding the analysis of the form.

The source material for this research consists of a selection of webcasts and televised broadcasts. The analysis examines transcribed verbal dialogue and subtitles as well as taking the full polysemiotic text into account. I have transcribed the dialogue word for word for all of the source texts. The subtitles for the web broadcasts were available as a transcript on their respective YouTube pages. I transcribed the subtitles for the DVDs. All of the data has been placed in tables juxtaposing the spoken dialogue with the corresponding subtitle. The comparison of transcribed dialogue and subtitle are based on Gideon Toury’s model of coupled pairs, where a segment of the TT is compared to the ST (Toury 1995, p. 89). Both the data extraction and data analysis follow the same model used in Pedersen (2011).

The investigation will examine both the form and content of the subtitles. The form will be investigated by extracting the following values from the subtitle transcript: number of lines per subtitle, condensation rate, subtitle density and expected reading speed. Since these values are based on Pedersen (2011), the formulas for the calculations mirror those in his work (ibid, p. 130):

1. **Expected reading speed**, defined as the speed with which the reader is expected to read a subtitle, as measured by the average number of characters per second of exposure time.

2. **Subtitle density**, defined as the number of subtitles the TT translation is divided into, as measured by the number of subtitles per minute. This measurement is based on subtitle quantity (i.e. the number of subtitles per TT), divided by the length of the TT in minutes.

3. **Condensation rate**, defined as a simple word count of the quantitative differences between ST and TT; presented as a reduction percentage from ST to TT. At this stage, no measurements of differences in information load or structure are made.

The expected reading speed is given by character per second, henceforth *cps*. The subtitle density value is given by subtitle per minute, henceforth *spm*. The condensation rate is presented as a percentage. Some of the source material included silences that lasted over 10 seconds, meaning that no verbal dialogue took place and therefore no subtitles were present for this time. These silences have been taken into account, they were excluded by timing them and subtracting the number of seconds from the total time and using this value to calculate the expected reading speed. These values will be included as footnotes and affect the calculations for expected reading speed.

The data will be analyzed by examining deviations from and/or use of subtitling norms. This will be measured by a close reading of both the transcribed dialogue (ST) and the subtitles (TT). Instances where subtitling norms indicate a certain behavior will be extracted and investigated in order to
determine whether or not they follow norms. It is not possible to study norms for dynamic media since the only norms that have been established are for static media. Therefore by pinpointing the differences and examining which approach is used instead it is possible to illustrate the nature of subtitled dynamic media without relying on existing norms.

The content will be analyzed by examining the subtitle transcript in relation to the original dialogue and the full polysemiotic text. The use or disregard of strategies within the translation theory field will be measured in part by using extralinguistic cultural references (ECR) as markers and examining how they are translated. An ECR is defined as “references to places, people, institutions, customs, food etc. that you may not know even if you know the language in question” (Pedersen 2011, p.44). These references have been extracted after a close reading of both transcripts. The strategy used for each ECR is determined and then analyzed in connection with the full polysemiotic text in order to conclude how this affects the viewer’s perception of the video. These strategies will also contribute to the conclusion of the skopos for each translation and whether a formal or dynamic equivalence is favored. In addition to ECRs, the content will be analyzed by locating omissions, additions, errors and deviations from subtitling norms.

This research does not aim at assigning a level of difficulty to the translation practice but it is necessary to point out the difference between a trained translator and an amateur. Untrained translators may work faster and more efficiently because they know less about complex theories. They have fewer doubts and do not waste time reflecting on the obvious. On the other hand, some awareness of different theories might be of practical benefit when confronting problems for which there are no established solutions, where significant creativity is required. (Pym 2010b p.4)

The difference is pointed out in order to analyze the data within a proper context, as Pym states above there are benefits on both sides of the equation. The difference between the final products from these two categories is similar to the difference between dynamic media and static media. Creativity and resourcefulness know no bounds and do not necessitate any formal training or funding which is also true for dynamic media. Professionals on the other hand benefit from similar conditions for static media, namely experience, resources, pre-existing reputations and industry connections. What an amateur lacks in knowledge can perhaps be made up for in creativity, and because they have never been taught the rules they can break or disregard them in order to achieve the final product they desire. Conventions, rules, and norms, however, incorporate fundamental benefits that are lost without them. The strict character limit, onscreen duration and other subtitling norms are followed so that the audience is always taken into consideration. It is easy to forgo these aspects in favor of a more natural or creative translation, but it is the audience that loses out if they for example do not have enough time to read the subtitle.
5. Analysis

This section includes a table of technical parameters; Table 1, as well as both individual analyses for each web broadcast followed by a general discussion. The individual analysis is made up of extracts that exemplify typical and/or deviant behavior. They are presented as coupled pairs, the first row contains the transcribed dialogue and the second contains the subtitle. If more than one subtitle is used in the example the rows always follow the order ST/TT/ST/TT. Each example is referred to by the name of the video followed by a time code. The individual analysis includes comparisons to the control group in order to demonstrate whether there was a deviation or compliance with subtitle norms. If such a comparison is not available, references to the norms in question are made. The general discussion compares the web broadcasts with each other and summarizes the findings from the full investigation.

5.1 Quantitative data

The compiled quantitative data is presented in Table 1 and is made up of quantitative values and calculated results. Four different values are assessed: condensation rate, subtitle density, expected reading speed and the number of lines in each subtitle. The videos are sorted by row into their assigned groups. The first two rows are made up of the control group, these values were placed first in order to clearly represent the conventional values that the remaining rows are compared to. The first column labels each video from the source material. The second, third and fourth columns presents values used to calculate the results presented in the fifth, sixth and seventh columns. The last three columns contain the number of lines in each subtitle.

The length of each video is included in order to illustrate that each video is of similar length. If the videos had been of drastically different lengths it could be argued that a longer video, for example a feature film, has more opportunities to deviate or comply with subtitling norms due to the increase in dialogue. This factor is also represented in the columns indicating the number of characters.

The full length is included for each videos, but some of them had pauses without verbal dialogue that lasted for over 10 seconds which were not included in the calculations for the condensation rate, expected reading speed and subtitle density. The values in the table represent the full length of the video, and the altered values that were used in the calculations are presented as footnotes.

The number of characters in the ST was obtained by transcribing the dialogue. The final number includes the dialogue verbatim as well as punctuation, spaces and dashes. This data is the most vulnerable since there is a certain margin of error for the transcript of the verbal dialogue as it was not possible to confirm by comparing it to an original manuscript. The next column displays the amount of characters in the TT that was directly obtained from the material. The subtitles for group 2, 3 and 4
were taken directly from their YouTube pages, the subtitle transcripts were available and I copied them into my own document, removed the time codes and counted the characters. I transcribed the subtitles for group 1 from the DVDs. The number of characters for both ST and TT are included in the table in order to visually indicate the difference in numbers as well as support for the calculations presented in the remaining columns.

The condensation rate was obtained using the formula presented in Pedersen (2011, p.130), the TT is subtracted from the ST and the difference is presented as a percentage. A plus or minus sign is added in order to indicate whether the value is a reduction or expansion. Expansion rather than reduction is an irregularity in this category and represents extreme deviations from subtitling norms.

The subtitle density was obtained using the formula presented in Pedersen (2011, p.130), the number of subtitles is divided by the length in minutes resulting in the number of subtitles per minute. The expected reading speed was also obtained using the formula presented in Pedersen (2011, p.130), the TT is divided by the length in seconds resulting in the amount of characters per second. Silences that last over 10 seconds have been omitted from the calculation of the expected reading speed. The values for subtitle density and expected reading speed were included in the table because they illustrate the average amount of subtitles the viewer has to read and the average amount of time that they are given to do so.

The last three columns specify the number of lines in each subtitle. They were included in the table to illustrate the difference in segmentation as well as a clear deviation from subtitling norms, which is using three or more lines.
Table 1: Quantitative data for all groups

Table 1 illustrates that the source material contains deviations from subtitling norms that increase within each group. This indicates a trend rather than solitary occurrences. It is also possible to see a variety within the deviations, not all webcasts deviate from the same norms. This is further exhibited in the more radical deviations, mainly found in Group 4. The text expansion evident in the positive percentage in three of the webcasts is unusual, but is even more so considering that only one of them also had a high cps and spm. This variety demonstrates that no specific subtitling norm is either disregarded or favored. This also indicates that whereas the subtitles for the televised broadcasts were supplied by a homogenous group, trained subtitlers, the subtitles for the webcasts are supplied by a heterogeneous group. Not only were most of the webcast subtitles written by amateurs, which is evident in the deviations, the variety between the deviations indicate that this a miscellaneous group of individuals.

---

\[1\] Value in calculation: 351 seconds. Reached by subtracting 73 seconds without dialogue, 424 – 73 = 351

\[2\] Value in calculation: 340 seconds. Reached by subtracting 38 seconds without dialogue, 378 – 38 = 340

\[3\] Value in calculation: 454 seconds. Reached by subtracting 156 seconds without dialogue, 610 – 156 = 454

\[4\] Value in calculation: 230 seconds. Reached by subtracting 70 seconds without dialogue, 300 – 70 = 230

\[5\] Value in calculation: 201 seconds. Reached by subtracting 199 seconds without dialogue, 400 – 199 = 201

\[6\] Value in calculation: 144 seconds. Reached by subtracting 23 seconds without dialogue, 167 – 23 = 144
5.2 Group analysis

5.2.1 Group 1: Professional televised broadcast/ professional subtitling

Each group in this section will be provided with the portion of Table 1 that is relevant to them. This has been done in order to make the references to the quantitative data directly accessible.

<table>
<thead>
<tr>
<th>Title</th>
<th>Length</th>
<th>Characters ST</th>
<th>Characters TT</th>
<th>Condensation rate</th>
<th>Subtitle Density/ smp</th>
<th>Expected Reading Speed/ cps</th>
<th>Single lines</th>
<th>Double lines</th>
<th>+Three Lines</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sandhamn</td>
<td>7:04</td>
<td>424</td>
<td>3485</td>
<td>-28%</td>
<td>-28%</td>
<td>9.2</td>
<td>7.1</td>
<td>30</td>
<td>35</td>
</tr>
<tr>
<td>HippHipp</td>
<td>4:00</td>
<td>240</td>
<td>3664</td>
<td>-24%</td>
<td>13.8</td>
<td>11.6</td>
<td>13</td>
<td>42</td>
<td>0</td>
</tr>
</tbody>
</table>

Table 1.1: Quantitative data for professional televised broadcast/ professional subtitling

This group includes the televised broadcasts Sandhamn and HippHipp. Although the clips are of different lengths the number of characters in the ST exhibits that there was a similar amount of dialogue. The story line in HippHipp is mainly derived through dialogue, whereas Sandhamn included more non-verbal action, which explains the lower cps.

The condensation rate did not deviate from subtitling norms. The text reduction included omission of repetitions and a reduction in content. The following example is taken from Sandhamn, a couple is arguing about a police investigation that the wife has become involved in.

*Example 1: Sandhamn 36:19*

ố, det är jävligt mycket som har hänt och det har påverkat oss allihop. Inte bara dig.
- Men det var ju jag som simmade på liket, det kanske är därför jag känner mig lite extra involverad.
- It’s affected all of us.
- I swam into the dead body.

This example shows that a considerable amount of the text has been reduced, the subtitles only contain the information that is relevant to the story progression, which follows subtitling norms.

Sandhamn and HippHipp are both closely tied to Sweden as a country and as a specific culture. Sandhamn takes place in Sandhamn, a specific area in the Stockholm Archipelago. The selection did not include any verbal references to culture, but visually it is very clear that the action takes place in Sweden and not in a nonspecific area. HippHipp on the other hand takes place in several Swedish cities and exaggerates different stereotypes. Keeping in mind that these subtitles are going against the current, as Gottlieb says, the subtitles have been adapted for an Anglophone audience which indicates a dynamic equivalence approach. HippHipp exemplifies this by both replacing Swedish references with English ones or paraphrasing the reference. Example 2 and 3 exhibit the use of substitution which is defined as “removing the ST ECR and replacing it with something else, either a different ECR from
the SC [source culture] or the TC [target culture] (Cultural Substitution), or something completely different that fits the situation (Situational Substitution)” (Pedersen 2011, p. 89). The following example is taken from a scene where a customer is negotiating with a store clerk over whether or not he should pay for a suit by claiming that he has had sponsors provide him with clothes before.

**Example 2: HippHipp 12:24**

<table>
<thead>
<tr>
<th>Jo, Erikshjälpen.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sure. Oxfam.</td>
</tr>
</tbody>
</table>

This is an example of cultural substitution, the ECR is the Swedish store name that has been replaced with the name of a store found in the UK. This occurs more than once in HippHipp, the only time a store or company is not substituted is when it is fictional and is then directly translated. The next example is taken from Sandhamn, a couple is fighting over a third person that has caused a conflict in their relationship:

**Example 3: Sandhamn 36:27**

<table>
<thead>
<tr>
<th>-Och det borde den där jävla dassdörren tala om för dig.</th>
</tr>
</thead>
<tbody>
<tr>
<td>-Dassdörr?</td>
</tr>
<tr>
<td>-And that barn door should tell you.</td>
</tr>
<tr>
<td>-Barn door?</td>
</tr>
</tbody>
</table>

The word “dassdörr” is translated to “barn door” instead of the direct equivalent which would be “outhouse door”. This is an example of situational substitution. Although many cultures have outhouses the Swedish term could also be referring to an indoor bathroom as well. The direct equivalent is longer than the replacement and disrupts the natural flow of the dialogue. By using “barn” instead the translator avoids specifying whether or not the bathroom is indoors or outdoors and uses a simple reference that is known in many different areas of the world. The character also goes on to explain what he means by this jibe which also applies to the replacement.

The following example illustrates a different strategy used to translate names. It is taken from a scene where a man is trying to sell a vacation package to a customer.

**Example 4: HippHipp 10:39**

<table>
<thead>
<tr>
<th>Holger, Glen Wish kommer han sen? Fikar du fortfarande?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Holger, is that Elvis guy coming?</td>
</tr>
<tr>
<td>Are you still having a break?</td>
</tr>
</tbody>
</table>

Glen Wish is an actual person and not a fictional character and although he may be known in Sweden he is not an international household name. This has been solved by replacing his name with a general description. This strategy is known as paraphrasing, when “this strategy is used, the ST ECR is
removed, but its sense or relevant connotations are kept by using a Paraphrase” (Pedersen 2007, p. 140). This strategy is used to achieve dynamic equivalence in the translation. Nida describes a dynamic translation to aim for “complete naturalness of expression, and tries to relate to the receptor to modes of behavior relevant within the context of his own culture” (Nida 1964, p. 144). By paraphrasing the cultural reference the TT attains a smooth and natural translation which does not contain any ambiguous or foreign information.

*Sandhamm* and *HippHipp* both use the verbal and visual channels in the subtitling which is apparent in the punctuation. Both videos include scenes where arguments take place and the characters yell and express their anger. The following example is taken from a scene where a man berates his employee for taking his meatballs from the office refrigerator.

**Example 5: HippHipp 10:48**

```
-Nej dom låg i allmänna lådan, där får alla ta köttbullar ifrån den där.
-But they’re my meatballs.
```

The viewer can see that the man is upset and hear that he is yelling so there is no need for the subtitle to use exclamation points. This can also be seen in example 3 where the word “dassdörr” is used, both characters are yelling when they deliver their lines.

The dynamic equivalence exhibited in the replacement of store names and the specific terms indicates that the skopos for both *Sandhamm* and *HippHipp* accommodates the target audience rather than preserving the ST. This also illustrates that the subtitler has taken advantage of the visual and auditory channels. They know that the viewer is aware of the Swedish setting and changes elements in the translation in order for the viewer to relate to and appreciate the content.

This group contains the only examples of static media and this analysis will be used in order to determine whether the remaining groups contain deviations or unconventional approaches. The analysis for Group 1 has exhibited different causes for text reduction, several translation strategies and examples of how the polysemiotic nature of the text is taken into account by the subtitler.

### 5.2.2 Group 2: Professional webcast/ professional subtitling

<table>
<thead>
<tr>
<th>Title</th>
<th>Length min/sec</th>
<th>Characters ST</th>
<th>Characters TT</th>
<th>Condensation rate</th>
<th>Subtitle Density/spm</th>
<th>Expected Reading Speed/ cps</th>
<th>Single lines</th>
<th>Double lines</th>
<th>+Three Lines</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Jaharâ</em></td>
<td>6:18</td>
<td>378</td>
<td>3226</td>
<td>2524</td>
<td>-22%</td>
<td>9.8</td>
<td>7.4</td>
<td>20</td>
<td>42</td>
</tr>
<tr>
<td><em>Rhodi</em></td>
<td>2:25</td>
<td>145</td>
<td>1212</td>
<td>1112</td>
<td>-8%</td>
<td>9.9</td>
<td>7.6</td>
<td>3</td>
<td>21</td>
</tr>
</tbody>
</table>

Table 1.2: Quantitative data for professional webcast/ professional subtitling
This group includes *Jaharà* and *Rohdi*. The videos have very little in common, the former is a scripted sketch and the latter an interview. Both are produced by an external company, *Jaharà* by a media production company and *Rohdi* by a museum. The expected reading speed and subtitle density did not deviate from the control group. The condensation rate for *Jaharà* was similar to the control group but *Rohdi* had a much lower value at -8%. This was not unusual because *Rohdi* is basically a monologue, the narration is given at a slower pace. There are also credits at the end of *Rohdi* that are subtitled, since these are not part of the verbal dialogue they were not counted in the ST. Both videos include text reduction, *Rohdi* has fewer instances and the subtitles are almost verbatim but some words have been omitted. The following example is the answer that the artist gives in regard to what the source to his inspiration is. The words that are not written in bold in the ST have been omitted from the TT because they are not important to the meaning of the message.

*Example 6: Rohdi 0:08*

<table>
<thead>
<tr>
<th>Oh ja, dels inte, ja <strong>ett ord nyfikenhet.</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>There is one word: curiosity.</td>
</tr>
</tbody>
</table>

*Jaharà* has higher text reduction, one of the characters often repeats the same word several times because he finds himself in an absurd situation and is struggling to remain calm. The following example takes place where two characters are reenacting a situation where a customer wants to return a product without a receipt:

*Example 7: Jaharà 1:36*

<table>
<thead>
<tr>
<th>Jaha, för du vet <strong>utan kvitto så kan vi på helheten inte göra någonting. Vi måste ha ett kvitto</strong> att utgå ifrån så att säga.</th>
</tr>
</thead>
<tbody>
<tr>
<td>We can't help you without a receipt.</td>
</tr>
<tr>
<td>We need to see one.</td>
</tr>
</tbody>
</table>

The text reduction follows subtitling norms, especially in *Jaharà* where longer segments have been omitted in the TT. *Rohdi* had less reduction but still followed norms. The lower value is explained by the subtitles not having to be shared by multiple speakers which allowed them to include more of the original dialogue.

A fictional store is referenced in *Jaharà* and is directly translated and a reference to a branch of the Swedish government uses the official equivalent. An official equivalent is defined as a term which requires “some sort of official decision by people in authority over an ECR” (Pedersen 2007, p.149). In this case the official decision has been made by the Swedish government who determine the English names of each branch. Both these examples lean towards more formal equivalence because they preserve the message from the ST, it is not adapted for a different culture. There is, however, an
example of dynamic equivalence from the same source material. The following example is from a shorter utterance given by one of the characters where he lists several obstacles he faces due to his physical disability:

*Example 8: Jaharå 1:47*

<table>
<thead>
<tr>
<th>Har aldrig sprungit i elljusspåret. Aldrig hoppat hopprep.</th>
</tr>
</thead>
<tbody>
<tr>
<td>-and never went for a run.</td>
</tr>
<tr>
<td>Never tried jumping rope.</td>
</tr>
</tbody>
</table>

The replacement of a specific term with a general description exhibits the use of the paraphrasing strategy. This occurrence indicates a dynamic approach and a focus on natural flow, it also indicates that the subtitler is relying on the visual channel. The viewer can see that the speaker is disabled, they do not need to know the specific route that the character cannot take it is enough that they know that one of his laments includes not being able to run. If this term had been translated directly the naturalness would have been disrupted and the character amount would have increased since the translation of “elljusspår” is “illuminated ski track” (Norstedts Engelska Ord 2014).

A deviation from subtitling norms occurs in the beginning of Rohdi. The only person shown on screen is a man and a woman’s voice is heard off screen asking him a question:

*Example 9: Rohdi 0:04*

<table>
<thead>
<tr>
<th>Vad har du för inspirationskällor?</th>
</tr>
</thead>
<tbody>
<tr>
<td>What are your sources of inspiration?</td>
</tr>
</tbody>
</table>

Subtitling norms require this utterance to be written in italics because it is heard off-screen, which the subtitler for this video has not done.

The subtitles for Jaharå mirror the content and production value behind it, the video is a comedic sketch that is both written and performed by two well-known figures in the Swedish media world. The skopos of the subtitles is to entertain a foreign audience by delivering the same quality to them as the audience for the original receives. Rohdi, while being a professional production, has not been created to the same standard as Jaharå which is also apparent in the subtitles. The video is much shorter and also unscripted, it is fundamentally a recording of a man answering one question. The purpose of this video is to inform interested parties instead of entertaining a wide audience. Although the subtitles have two different skopoi, both follow subtitling norms and have relatively similar values.

This group does not exhibit the same level of dynamic equivalence as Group 1, the condensation rate is lower but the text reduction still shows that the translator has chosen a more conventional translation rather than a word for word transcript. However, the paraphrasing strategy speaks for an intent to achieve a natural flow so the final result still leans towards dynamic equivalence rather than formal.
Subtitling norms have been followed by reducing the text in order to stay within the character limit as well as synchronizing and displaying them at a speed that the viewer can read. These subtitles could be used for static media, however, the deviation found in Rohdi where the subtitle for an off-screen speaker was not written in italics indicates that there is a difference between subtitles for static and dynamic media.

5.2.3 Group 3: Independent professional webcast/ amateur subtitling

<table>
<thead>
<tr>
<th>Title</th>
<th>Length min/sec</th>
<th>Characters ST</th>
<th>Characters TT</th>
<th>Condensation rate</th>
<th>Subtitle Density/spm</th>
<th>Expected Reading Speed/ cps</th>
<th>Single lines</th>
<th>Double lines</th>
<th>+Three Lines</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sanning</td>
<td>10:10</td>
<td>610</td>
<td>2959</td>
<td>-5%</td>
<td>11.4</td>
<td>6.1</td>
<td>91</td>
<td>23</td>
<td>2</td>
</tr>
<tr>
<td>BOB</td>
<td>5:00</td>
<td>300</td>
<td>1571</td>
<td>+5%</td>
<td>15</td>
<td>7.2</td>
<td>73</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>Sneakerliv</td>
<td>8:45</td>
<td>525</td>
<td>6069</td>
<td>-13%</td>
<td>12.6</td>
<td>10.1</td>
<td>48</td>
<td>62</td>
<td>0</td>
</tr>
</tbody>
</table>

Table 1.3: Quantitative data for independent professional webcast/ amateur subtitling

The videos in this group are produced by independent filmmakers, which means that the creators have a certain knowledge of the medium but the production is not the result of a commission from an external source. The group includes Sanning, BOB and Sneakerliv. While the subtitle density and expected reading speed were similar to the control group, however, the condensation dropped considerably.

The text reduction for this group was mainly low, but in BOB there was expansion which is indicated in the table by a plus sign. The verbal dialogue in BOB was subtitled almost verbatim, this coupled with the subtitling of non-verbal dialogue caused the text expansion. Text expansion is very unusual in subtitling and therefore considered an extreme deviation. Sanning displayed text reduction but the values are still quite low, the highest level was found in Sneakerliv. Repetitions and superfluous filler words were removed in both cases. However, Sneakerliv included instances of proper names being replaced by abbreviations. The following example is taken from a story about how the speaker first took an interest in sneakers.

Example 10: Sneakerliv 0:55

| När jag var 12 eller 13 år så kom min bror hem med ett par Air Jordan 5. |
| When I was 12 or 13 my brother came home with a pair of AJ5. |

This abbreviation shows two things: the first is that the subtitler has taken the visual channel into account because a pair of Air Jordan 5 sneakers is shown directly after this is said. The second factor is that the subtitler has a specific target audience in mind. The documentary itself is quite niched, it
focuses very narrowly on sneaker collectors, their collections and the history behind their collection. There is very little focus on any aspect outside of this realm, no connections are made to other areas of interest. There is a regular use of sneaker-related terms used continuously throughout the dialogue. This would allow the subtitler to assume that the viewer is already interested in this topic and therefore familiar with specific terms as well as abbreviations. In this light, it is possible to assume that the subtitles for this video lean towards formal equivalence because it is the information that the audience is interested in. This is reinforced by the genre, *Sneakerliv* is a documentary and therefore an informative video. However, this group has illustrated that genre does not necessarily dictate the skopos, *BOB* also displayed formal equivalence by preserving both form and content in the ST, even though it is a fictive story with the purpose of entertaining.

The formal equivalence is evident in the low text reduction and direct translations, no strategies such as paraphrasing or cultural substitution were used. When the subtitles give an almost verbatim rendition of the dialogue, both the form and content of the ST are closely preserved, which complies with formal equivalence which Nida states “is concerned with such correspondence as poetry to poetry, sentence to sentence, and concept to concept” (Nida 1964, p. 144). The subtitles that exhibit formal equivalence match the word order closely except when there is a conflict with the target language grammar, matching both the form and content of the ST.

One distinct deviation from subtitling norms was the occurrence of a subtitle with three lines in *Sanning*. The following example occurs when two characters meet for the first time and one of them wants the other to leave.

**Example 11: Sanning 1:03**

<table>
<thead>
<tr>
<th>Du får dra någon annanstans, den här platsen är upptagen.</th>
</tr>
</thead>
<tbody>
<tr>
<td>You’ll have to find another spot, this one is taken.</td>
</tr>
</tbody>
</table>

The subtitle has 52 characters, which is well below the limit for two lines. There are no technical reasons for this to be split into three lines and the content does not indicate any reasons either. There is a second example of more than two lines. This example is taken from a scene where the characters are sitting on a cliff playing a game, one of them has told a lie and now has to toss her shoes over the edge.
Example 12: Sanning 4:00

- Båda skorna,  
- Räknas som en, jag vet.  
- Jag ville vara fin när jag...  
- Both shoes,  
- Count as one, I know.

57  
00:04:09,000 --> 00:04:11,200
I wanted to look good when I was...

The inclusion of the numbers suggest that this subtitle with five lines was not intentionally divided in this way like the previous example with three lines. The last sentence appears to be a separate subtitle and the numbers are assumed to be the time code for it. However, this error was left in the final product, which means that it will be analyzed as a five line subtitle for the purpose of this research.

The use of dashes to indicate speakers follows norms for most of the material but there are a few exceptions. In Sanning the same character utters two single word sentences, one directly after the other. However, the subtitle is split into two lines with a dash before each word. The following example occurs when a character is speaking to her mother on the phone:

Example 13: Sanning 0:30

Hallå? Mamma?
- Hello?  
- Mum?

Subtitling norms would not require the use of dashes for this case and would not divide the subtitle into two lines. This is a singular occurrence; the rest of Sanning uses dashes according to norms.

This group exhibited deviations from norms regarding speakers heard off screen. When a character speaks to her mother on the phone the mother’s voice is heard off screen. The viewer knows who the speaker is from the previous statement made by the girl where she says her mother’s name into the phone which she is still holding to her ear as the second voice is heard. Subtitling norms require this subtitle to be written in italics.

Example 14: Sanning 0:33

Du vet väl att du inte kan ringa mig på jobbet?  
You know you can't call me at work.

Off-screen sounds are also present in BOB. These are also not written in italics but the subtitler has indicated that they were different from the rest by putting them in parentheses, example 15: “(Meow)”
(BOB 0:11) and example 16: “(Scream)” (BOB 2:41). Both sounds are relevant to the plot and the on-screen character reacts to them.

This is not only a deviation from subtitling norms, but the inclusion of these sounds would only apply to SDH. A viewer of interlingual subtitles can hear both of these sounds and do not need them to be translated. Their inclusion is also one of the reasons for the text expansion, which is also a deviation from subtitling norms.

Group 3 has exhibited several deviations from subtitling norms which are evident in the quantitative data. The condensation rate is below average and considerably lower than Group 1 and the text expansion illustrates the first major deviation from subtitling norms. The alternative approaches to form that started to appear in Group 2 continue here, appearing in the punctuation and subtitle segmentation. These deviations also reflect the flexible nature of dynamic media, the errors that are left in the final product are not permanent, they can be amended without removing the original and replacing it with a new version.

5.2.4 Group 4: Amateur webcast/ amateur subtitling

<table>
<thead>
<tr>
<th>Title</th>
<th>Length min/sec</th>
<th>Characters ST</th>
<th>Characters TT</th>
<th>Condensation rate</th>
<th>Subtitle Density/ spm</th>
<th>Expected Reading Speed/ cps</th>
<th>Single lines</th>
<th>Double lines</th>
<th>+Three Lines</th>
</tr>
</thead>
<tbody>
<tr>
<td>vlog</td>
<td>2:51</td>
<td>171</td>
<td>1868</td>
<td>-9%</td>
<td>19.3</td>
<td>9.9</td>
<td>52</td>
<td>3</td>
<td>0</td>
</tr>
<tr>
<td>Gathering</td>
<td>6:40</td>
<td>400</td>
<td>2178</td>
<td>-26%</td>
<td>4.5</td>
<td>8</td>
<td>14</td>
<td>16</td>
<td>0</td>
</tr>
<tr>
<td>Mealtime</td>
<td>2:47</td>
<td>167</td>
<td>496</td>
<td>+103%</td>
<td>6.1</td>
<td>7</td>
<td>0</td>
<td>15</td>
<td>2</td>
</tr>
<tr>
<td>Haffa</td>
<td>2:44</td>
<td>164</td>
<td>2210</td>
<td>+15%</td>
<td>28.6</td>
<td>15.5</td>
<td>71</td>
<td>7</td>
<td>0</td>
</tr>
</tbody>
</table>

Table 1.4: Quantitative data for amateur webcast/ amateur subtitling

This group is made up of videos that are both created and subtitled by amateurs, meaning they have created the videos as a side project rather than profession and it is assumed that they have no formal training within either disciplines.

The condensation, expected reading speed and subtitle density varied within the group as well as from the rest of the source material. The condensation rate for two videos, Mealtime and Haffa, were positive which indicates that there was an expansion instead of a reduction. The addition of text is highly unusual and is explained in both cases by extreme deviations from subtitling norms.

The only webcast in this group that had displayed values similar to the control group was Gathering. The condensation rate was -26% which is slightly below average but closer than most of the other source texts. The text reduction followed subtitling norms, omitting information that was not necessary for the viewer to understand. The following example is from a montage, a person on stage is
speaking to a crowd and it is assumed that he is both accepting praise as well as explaining the content of the videos he creates for YouTube:

Example 17: Gathering 0:34

-Yes, well.
Thanks. I try to do comedic things.

This subtitle omits one speaker entirely and reduces most of what the second speaker says but conserves the core of his statement. The form and content of the ST have been altered which goes against formal equivalence. Only the message has been preserved, the tone from the ST is not evident in the TT, which is much more impersonal. This change in tone speaks against dynamic equivalence, however, the TT has still preserved the message which tips the scales in favor of it.

At the other end of the spectrum, the expansion in Haffa can be explained by the labels that identify the gender of each speaker. Each subtitle begins with a label that reads either “Guy”, “Girl” and in one case “Father”. The following example depicts a young man approaching two strangers and asking if they are involved in a relationship:

Example 18: Haffa 1:50

<table>
<thead>
<tr>
<th>År ni tillsammans?</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Guy</strong> - Are you guys in a relationship?</td>
</tr>
<tr>
<td>Nej.</td>
</tr>
<tr>
<td><strong>Girl</strong> - No</td>
</tr>
</tbody>
</table>

There is no technical reason to add these labels, the viewer can both see who the speaker is and what their gender is. They can also hear the voices clearly which also signal the gender of the speaker. Because of the content of the video, a young man flirting with unknown women in a public area, it is possible to deduce that the reason no proper names are used is because the identities of the women are unknown. The young man is part of the group that created the video and is perhaps labeled “Guy” in order to remain anonymous. Regardless of what the labels are, their presence is highly irregular.

The expansion in Mealtime is due to several subtitles consisting of what appears to be HTML coding and not a translation of the dialogue. The following example is from a segment that illustrates that potatoes are needed in the described recipe:

Example 19: Mealtime 0:27

Potatisar i plural.
{vtf\ansi\ansicpg1252\cocoartf1187\cocoaubrtf370\fonttbl f0\fswiss fcharset0 Helvetica;}
This is clearly not a translation but a technical error that has not been corrected. This is not the only error in this video, almost all of the subtitles are unsynchronized with the audio. When the subtitles do translate the dialogue they appear a few seconds after they have been said and some lines are delivered after the subtitle has disappeared from the screen. Another error that is also a deviation from subtitling norms is the use of slashes after each subtitle line. The following example is from a segment that illustrates that mushrooms and carrots are needed in the recipe.

*Example 20: Mealtime 0:36*

<table>
<thead>
<tr>
<th>Svamp. En morot.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mushrooms\</td>
</tr>
<tr>
<td>One carrot\</td>
</tr>
</tbody>
</table>

Every line except those that only consist of HTML coding end with a slash. This also seems to be a technical error that has remained uncorrected. There is also an error in the translation. The following example shows the final lines in the video where the prepared meal is eaten by the main character and his guests.

*Example 21: Mealtime 2:28*

<table>
<thead>
<tr>
<th>Swedish red beets, they're good for you.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nästa gång äter vi hans skägg.</td>
</tr>
<tr>
<td>Swedish redbreast, they're good for you\</td>
</tr>
<tr>
<td>Next time we eat his beard\</td>
</tr>
</tbody>
</table>

This subtitle exhibits three different deviations from norms; a faulty translation, it is unsynchronized with the verbal dialogue and there is visible coding. The translation of “red beets” to “redbreast” is incorrect, a redbreast is a species of bird and not another term for beets. This line was also delivered in English so there is no reason for the subtitle to use a different word which points to this being a spelling error. The line “Nästa gång äter vi hans skägg” is delivered 8 seconds after the subtitle has disappeared from the screen. The slash and bracket are assumed to be visible coding.

Other mistakes within this group are found in vlog which includes spelling errors. The following example is from a description the speaker gives about how she is learning to speak Swedish:

*Example 22: vlog 0:26*

<table>
<thead>
<tr>
<th>Och jag övar på varje sätt som jag kan hitta på.</th>
</tr>
</thead>
<tbody>
<tr>
<td>And I practise in every way that I can think of.</td>
</tr>
</tbody>
</table>

The speaker informs the audience that the purpose of this video is for her to practice her Swedish, which indicates that it is not her first language and that she has not yet mastered it fluently. She introduces herself as a Russian native which also indicates that English is not her first language either.
The content for both dialogue and subtitle being two different foreign languages for the user can explain any spelling or grammatical errors, it is unlikely that they appear the final product due to negligence.

The previous examples from *Haffa* and *Mealtime* also illustrates improper use of punctuation. The only punctuation used is question marks in *Haffa* and commas in *Mealtime*, there are no periods to indicate the end of a sentence. Dashes are used to label the speaker, not to indicate two speakers in the same subtitle. There is also a notable example of misused punctuation in *vlog* where parentheses make an appearance. They are, however, not used in the same way as they are in *BOB*, in this case they are used to indicate a side comment. The following example is part of the speaker’s assessment of her own level of fluency of the Swedish language:

*Example 23: vlog 0:35*

<table>
<thead>
<tr>
<th>Aven om jag gör fel.</th>
<th>even if I make mistakes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Det gör jag förstås.</td>
<td><em>(and of course I do)</em></td>
</tr>
</tbody>
</table>

The parenthesis mirrors the change of tone in her voice. The utterance within parenthesis is an admission that is not directly related to her full statement but additional information. The full monologue is similar to a blog or diary entry, if this text had appeared as written text the use of parentheses in this instance would not be out place.

Parentheses were used in three different ways in three different videos: to indicate a side comment, to indicate an off-screen speaker and to explain a culture specific term. This wide variety indicates that webcast subtitlers are a diverse group who have appropriated the creative freedom that written text as on the Internet to include subtitles as well. Whether the strategy itself is necessary is a separate issue. Since the audience can hear and see the action corresponding to the punctuation, the parentheses are not necessary. Parenthesis are conventionally used in subtitling to “set apart relevant but supplementary information that could be dropped without changing the meaning of the sentence” (Díaz Cintas & Remael 2007, p.108). However, they are very rarely used and subtitlers often try to incorporate the material with commas instead if they are not omitted completely (ibid). The only case that this would apply to is *vlog*, and then only to a certain extent since the line within parenthesis could also work within a comma. The use of parenthesis, a rare occurrence in static media subtitles, appearing in a third of the webcasts indicates a clear difference between subtitling for both groups.

Another factor related to punctuation is the use of dashes to indicate a new speaker. None of the webcasts in this group have used them for this function. In *vlog* there is only one speaker and no dashes would be required. *Haffa* uses almost exclusively single lines and has both a very high expected reading speed and subtitle density, each new speaker is given their own subtitle. *Mealtime*
has only one instance of two speakers and the lines are merely separated into one subtitle with two lines. *Gathering*, however, places two speakers in the same subtitle in the same line several times. The following example is from a person off camera asking a young man on screen a question:

**Example 24: Gathering 2:44**

| -Viktor hur gammal är du?  
| -Tjugoett.  
| How old are you Victor? I'm twenty-one. |

Both speakers are placed in the same line, relying on the viewer to distinguish between the difference voices and work out who says what for themselves.

There are culture specific references in all of the videos which are dealt with in different ways. Haffa translates the slang word for an area in central Stockholm by adding explanatory information. This strategy is known as specification which “means retaining the ECR in its untranslated form, but adding information that is not present in the ST, making the TT ECR more specific than the ST ECR” (Pedersen 2007, p.131). The following example involves the young man attempting to seduce a stranger, they are located at the referenced location.

**Example 25: Haffa 2:26**

| Har ni kysst någon någon på Plattan en gång?  
| Guy - Have you kissed someone at Plättan (place in Stockholm) before? |

The subtitles for Haffa have exhibited formal equivalence by mirroring the form and word order as well as translating the content almost verbatim. This strategy indicates that despite this, the skopos is still to engage the viewer by providing as much information as possible for them to understand the what is being said. The only instance of text reduction is apparent in this example, the speaker repeats the word “någon” which is omitted in the subtitle.

**Mealtime** is part of a webseries about Swedish cooking, so the references in this webcast are tied to food. The video revolves around the recipe a typical Swedish dish, “pyttipanna”. This dish is not translated in the title “Chop Chop Carnage Stew (Alternative Pyttipanna)”. This appears to be a sort of paraphrase coupled with specification.

**Gathering** deals the brand of a specific Swedish beer by using paraphrases. The following example is a scene where several people are sitting in a group drinking beer, the camera pans over the different participants as the narrator names them.

**Example 26: Gathering 3:14**

| Michael, Alexandra and Amanda drink light beer. They drink northern beer. |
The name “Norrlands” is part of the full name “Norrlands öl” which is a brand name. There is no official equivalent and instead of keeping the original name the subtitler references is by preserving one defining characteristic, namely the location it comes from. This is another example of dynamic equivalence, if the original brand name had been used in the subtitle this could alienate the viewer by reminding them that they are not part of the culture that they are observing.

One deviation from subtitling norms that appears throughout Gathering in different forms is the subtitling of songs. Subtitling guidelines state if a song is relevant to the story, i.e. not part of the soundtrack playing in the background, the lyrics should be written in italics (Díaz Cintas & Remael 2007, p. 127). The following example is from a scene where several people sang “Ja må han leva”, the Swedish song that is sung on birthdays.

Example 27: Gathering 1:13

<table>
<thead>
<tr>
<th>I hundrade år.</th>
<th><strong>birthday song</strong></th>
</tr>
</thead>
</table>

The subtitle did not translate the lyrics but instead contained a title for the English version in quotation marks. Another example of this occurs when a woman performed on a stage in front of a live audience, the lyrics of the song were not translated but the subtitle instead indicated the identity of the performer, example 28: ““Melanie Wehbe performs”” (Gathering 1:42).

A different performance was subtitled as commentary rather than translation or transcription, the scene shows a man performing a song on stage in front of a live audience:

Example 29: Gathering 2:03

| But in the world of corporation there is no guarantee | **“First ever live version of Ogooglebar”** |

The subtitle for this utterance did not need to translate it as it is in English, but instead named the song that was performed and included additional information regarding the context that was not included in the ST. The viewer watching without subtitles did not receive any indication that this was the first time that this song was performed which means that the TT to contains more information relevant to the story than the ST. The inclusion of this additional information goes against subtitling norms since it is not part of the spoken dialogue.

These two examples indicate a dynamic equivalence that transcends the linguistic aspect. There are many quick cuts and edits because the video is a montage of scenes from an event so there is not always enough time for the viewer to fully absorb what they are watching. The subtitles make up for this by being short and informative. Example 28 informs the audience of the singer’s identity rather
than translating a portion of the lyrics. The lyrics in this case are not important, it is the performance that is the focus. Example 29 also ignores the lyrics, instead they inform the viewer of which song it was with the additional information that this is the first performance of this particular song which was deemed more important than the lyrics, which are in English but difficult to make out.

Finally there were two instances of a song where the lyrics were neither translated or transcribed, the subtitles indicated the titles of the songs being sung instead. Both of the below examples depict a group of people singing in a park:

*Example 30: Gathering 4:00*

| Oh whoopdedoo, jag vill va som du |
| "King Louie song from Djungle Book in Swedish" |

*Example 31: Gathering 4:40*

| I just have sex with boys. |
| "IJustWantToBeCool and RoomieOfficial" |

The first example includes an example of specification and the second appears to be titles from two different songs which possibly indicates that it is actually a medley. The form and content of the ST have been changed completely which indicates that formal equivalence has not been imposed. The audience is supplied with extra information to make up for the references specific to Swedish culture in order for the target audience to connect to the TT as much as possible. It is this factor that speaks for dynamic equivalence, even though the tone from the ST is lost by titling or describing the songs rather than translating the lyrics. The dynamic equivalence reveals the skopos to be to entertain the audience as much as possible. By providing them with relevant information instead of a verbatim translation, the subtitles engage the audience of the TT on the same level as the audience of the ST.

*Mealtime, Haffa* and *vlog* all displayed preferences for formal equivalence and even though the condensation rates were both negative and positive the text reduction was minimal. However, *Gathering* had a clear preference for dynamic equivalence. *Gathering* is not only a video that is available on YouTube, but it is also about YouTube videos and their creators. It consists of a montage of clips from a ceremony celebrating several YouTube filmmakers as well as a party afterwards. The dynamic equivalence exhibited in the subtitles indicates that the skopos is to connect to and engage as many viewers as possible which is very much in the spirit of the videos that *Gathering* celebrates.

This group has exhibited the most extreme and frequent deviations from subtitling norms out of all of the webcasts. These subtitles would not be acceptable for static media not only due to the disregard of norms but also because of the form and content. Their form exhibited technical errors and subtitles with more than two lines. The content included incorrect translations, additional irrelevant information and the strategies reflect those used to translate written text and not polysemiotic media. These factors
illustrate how far text for dynamic media can digress from both subtitling norms as well standard rules for grammar, punctuation and spelling.

5.3 General discussion

The analysis of all of the webcasts has exhibited that each group moves further and further away from the subtitling norms found in their static media counterparts. For televised broadcasts, “what counts more than anything else is the transfer of speech acts, not necessarily the exact elements that make up the original speech acts” (Gottlieb 2008, p.3). The analysis has shown that this claim is not true for webcasts where the majority had a low condensation rate and many were subtitled almost verbatim. The analysis has also demonstrated a preference for formal equivalence in the majority of the webcasts. These factors allow for a deduction of the skopos to be made which reflects an informative function despite the fictive content for most of the webcasts.

The individual analysis of each group presented the following results. Group 1 acted as the control group and exhibited the standards and subtitling norms that were expected from static media. Group 2 began to show tendencies towards alternative solutions. Even so, both the contents of the webcasts and the subtitles were not exclusive for dynamic media, but could have been broadcast on television as well. For Group 3 the divide between static and dynamic media is more established. The most apparent difference was found in the minimal text reduction, but there were also errors that would be unacceptable for static media. The final group, however, contained the most extreme approaches. These include text expansion instead of reduction, technical errors and spelling errors.

Looking at the groups in the order in which they are presented illustrate that there is a trend moving in one direction rather than sporadic deviations. This phenomenon is clear in the form and presentation of the subtitles, which gradually differs from the control group. Each group takes the form one step further resulting in a wide gap between Groups 1 and 4. This gap establishes a clear difference between subtitles for static and dynamic media. This difference in form indicates that the norms in place for televised broadcasts are not applied to webcasts, which results in subtitles for dynamic media having a different visual appearance than those for static media.

The webcast subtitles have also exhibited a difference in content. The majority of the webcasts translated the dialogue very literally. The majority favored formal equivalence, which resulted in considerably less text reduction than the televised broadcasts. The content in the subtitled webcasts included almost all of the information communicated in the verbal dialogue. The televised broadcasts on the other hand both omitted information and replaced cultural references with equivalents from the target culture. This difference determines that the skopos for the webcasts is to communicate as much of the information in the ST as possible. The televised broadcasts’ skopos is for the audience of the TT
be entertained to the same standards as the audience of the ST. This difference in skopos is one of the factors that separate subtitling for dynamic media and static media.

Another revealing factor found in the content was the inclusion of errors. Surprisingly the video with the most views, *Mealtime*, also had the most errors out of all the subtitles. The video itself is edited and assumed to be scripted which indicates that a certain amount of time and work was put into it. The subtitles were almost always out of synch, all of them included a slash to indicate a line break and some of them were entirely written in code. There were also examples of coding present in the subtitles for *Sanning* which resulted in a subtitle with five lines. These errors are technical issues, they are not deviations from subtitling norms or incorrect translations. Their inclusion in the final product separates them from televised broadcasts in two ways. First and foremost they demonstrate that webcast subtitles have fewer obstacles on their way to the finish line. Much like texts that are published online that do not need an editor to proofread them before they are published, these errors prove that this is also true for subtitles. The second aspect is a reminder that webcast subtitles are not permanent; they can be removed or altered by the user at any time. Subtitles for televised broadcasts, DVDs and cinematic screenings cannot be altered in the same way. For example a DVD with errors in the subtitles would have to recall all the physical copies, destroy them and produce new ones in order to amend the problem. A webcast can change their subtitles with a few clicks of a button; they can be altered instead of replaced. It is possible that it is this that allows for errors, the subtitler knows that they can subsequently fix any problems, so they are not as strict during the initial process. These two factors are the same parameters used to separate dynamic media from static media so it is not unexpected that the subtitles in both fields reflect the same differences.

Another possible reason for the difference in content in the webcast subtitles is the disregard for the visual and audio channels that create a video, which caused many of the deviations from subtitling norms. The minimal text reduction speaks for a formal equivalence and a more literal approach to translation. This was, however, not a systematic occurrence. Both *Jaharâ* and *Gathering* exhibited text reduction that included more than solitary words and only the former was professionally translated. However, the majority ignored the polysemiotic nature of the translation and relied on the verbal dialogue alone. To a person who has never translated text for any kind of interactive media, relying on audio and visual channels is not an obvious choice. Their goal is to translate the verbal dialogue into another language and present it as subtitles. The majority of the source material indicates that this results in an almost verbatim translation and a preference for formal equivalence. This creates a contrast with the control group that displayed a preference for dynamic equivalence, which is also what most televised broadcasts favor (Pedersen 2007, p. 7).

It is possible that this disregard is caused by the subtitlers’ lack of education and training within both AVT and translation studies in general. This is not unexpected for the webcast subtitles considering that amateurs wrote all but two of them which follows norms. Another reason for a lack of education
is that since webcasts are a relatively new form of media there has not been much research within their field yet. The only norms and standards for subtitling that exist today are for static media and it would therefore not be possible for anyone to receive formal training as a webcast subtitler. A similar field where subtitles have not been fully researched is video games. O’Hagan & Mangiron describe subtitling for video games as lacking standardization and of poor quality. Several of the issues that they bring up are also present in the source material for this thesis, including three or more lines for each subtitle, extreme variation in characters per subtitle and short exposure time (O’Hagan & Mangiron 2013, p.290). They attribute these deviations to be “in part due to a low level of awareness, exacerbated by a general absence of AVT training among game localizers along with cost implications” (O’Hagan & Mangiron 2013, p.290). This is also true for subtitles for dynamic media, there are no specific norms for this field and therefore no training can be provided. Professionals who have already worked with static media treat subtitles for dynamic media in the same way that they would a televised broadcast. The amateurs working with them have no subtitling or localization training and therefore treat them like any other written text they would release online.

The source material has demonstrated that there is a different power balance for dynamic media than there is for static media. The majority of the subtitles were written by someone involved in creating the original content. It is beyond the scope of this research to interview each subtitler in order to investigate their exact qualifications, but it is assumed that their only connection to subtitles is from their own experience of watching subtitled media. The shift from audience to creator in the webcast subtitles illustrates the difference in boundaries for dynamic and static media. The audience may have influence on the creative process for static media, but they do not play an active role. This shift also explains why webcast subtitling does not follow normative behavior and have a considerably different appearance than subtitles for televised broadcasts. The subtitlers that write for dynamic media are a separate group than those writing for static media. The platform has made it possible for this group to exist, and the content has given them a reason to exist. But it is the group itself that causes the subtitles to appear in the way that they do.

It is necessary to point out that the focus on deviations and disregard of norms is not a form of negative criticism. Technically, rules have been broken but the source material has not necessarily suffered from it. On the contrary, dynamic media has been depicted as an eclectic, individual and egalitarian field. The analysis has provided examples of both professionals following norms and amateurs making up new ones. Which is precisely how norms are created in the first place. “In language, as in human behavior in general, motivated and repeated violations of norms, undertaken by persons – in case subtitlers, dubbers and journalists – who communicate with mass audiences, eventually become new norms” (Gottlieb 2001b, p. 258). Perhaps the deviations described in this thesis will one day become common practice within the field. However, today they belong to the minority and only time will tell how long they stay there.
6. Conclusion

This thesis has examined what subtitles for webcasts look like today by examining their form and content. Since no previous research has been done on this subject, it was necessary to ensure that the analysis would treat webcasts and televised broadcasts as two separate entities. This required the construction of two categories, static media and dynamic media. By separating the two types of media from each other it was possible to analyze each side on its own and then compare their differences.

The analysis showed that there is a clear difference between subtitles for the two types of media in several areas. The content differed mainly because of the translation strategies and approaches used by the subtitler. The majority of the webcasts displayed formal equivalence within their translations, delivering subtitles that were almost word-for-word translations of the dialogue. The televised broadcasts displayed the opposite, favoring dynamic equivalence and providing the foreign audience with subtitles that relate to their own cultural references.

The form varied between dynamic and static media, but the webcasts were also very different from each other. They showed several approaches to the presentation of subtitles, both positive and negative. Most were inventive, but there was also a certain level of neglect that was evident in the inclusion of errors. The main result that the deviations in both form and content have shown is that there is a very diverse range of subtitling being done on the Internet today. The source of these deviations has not been attributed to the content of the videos. The selection included both fiction and non-fiction, professional and amateur. There were similarities between completely different types of media and differences within the same genre.

The platform, however, did affect the subtitles but was not solely responsible for the final results. Instead the available technology and digital environment functioned as advantages rather than determining factors. The main benefit that dynamic media holds is the ability to integrate and connect. One component can connect to another through a link, and one culture can connect to another just as easily. Translation theory, and in this case AVT especially, can help this connection by overcoming the language boundary. If webcasts continue to evolve and grow into one of the main forms of media distribution, it is necessary for the translation within this field to develop as well. Before webcasts, DVD was the new kid on the block, taking the practice into the digital age by adding new functions such as optional subtitles and allowing the audience to choose between languages. Webcasts are the next step that takes subtitling to another level, which includes providing the audience with more language choices, the ability to customize subtitles and even invites them to partake in the creative process.

Audience involvement was evident in the subtitles examined in this thesis, which included contributions from people outside of the translation community as well as professionals. The freedom
that exists on the Internet allows for unlimited contributors and fewer restrictions. The effect this has on subtitling is that it cuts out the middleman. The only cases where the subtitles were written by a person who was not involved in the creation of the original video were the professional webcasts. The examined deviations illustrate the different approaches that each subtitler has taken when dealing with a new form of media. Those who have worked with subtitles for static media treat subtitles for dynamic media in the same way, which involves following norms. Those that are new to subtitles treat them as they would any other dynamic text, which involves no required editing or proofreading from an external source. This means that the form and content of these subtitles are closely tied to the person writing them. They do not share the same boundaries that professionals do and can present whatever they want. This does not mean that professionals produce completely homogenous subtitles without any idiosyncrasies, but since they follow norms they are still relatively uniform. Therefore, the difference between subtitling for televised broadcasts and webcasts has been attributed to the creator of the subtitles, not the content or the platform.

This thesis has not proposed any new norms, solutions or strategies. Instead, it has presented a portion of the subtitling landscape on YouTube and what it looks like today. That landscape can be likened to a teenager, wild and curious, testing boundaries and searching for a way to express itself. In order for it to grow it needs guidance from those that are older and wiser, who once upon a time were teenagers themselves but have developed into the adults they are today. Subtitling for webcasts can definitely benefit from norms that have been established for their televised counterparts but it is also necessary for them to create some of their own.
Primary Sources


Hedefors Productions. 2013. BOB (Kortfilm/Short Film) ENG-sub available! [online video] Available at: <http://www.youtube.com/watch?v=gXzfiJUp9s> [Accessed 12 March 2014]


Morden i Sandhamn. 2010. DVD, Sweden: Filmlance International AB. Directed by Marcus Olsson


Stratos Cinema. 2011. Sanning eller Död (Kortfilm / Short Film) [video online] Available at: <https://www.youtube.com/watch?v=VCDcZ9xNiwc> [Accessed 12 March 2014]

Secondary Sources

Bibliography


Burgess, Jean and Joshua Green. 2009. YouTube: online video and participatory culture. Cambridge: Malden


**Filmography**


*Dr. Seuss’ The Cat in the Hat.* 2003. Motion picture. USA: Universal Pictures. Directed by Bo Welch


*Punk’d.* 2003. Television broadcast. USA: MTV. Directed by Brendan Carter


*Vice.* 2013. Television broadcast. USA: Vice Media. Produced by Johan Kaplan