The Efficacy of Subtitles
Three Different Subtitling Conditions to Enhance FL Vocabulary Knowledge

Monique Bos
s1970143

MA in Applied Linguistics
Faculty of Liberal Arts
University of Groningen

Supervisor:
Dr. H. Loerts

Second reader:
Prof. dr. C. L. J. de Bot

July 7th, 2014
Word count: 16,006
# Table of Contents

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abstract</td>
<td>4</td>
</tr>
<tr>
<td>1. Introduction</td>
<td>4</td>
</tr>
<tr>
<td>2. Background</td>
<td>7</td>
</tr>
<tr>
<td>2.1 Film in the classroom</td>
<td>7</td>
</tr>
<tr>
<td>2.2 Overcoming language boundaries</td>
<td>9</td>
</tr>
<tr>
<td>2.3 Multimodal processing</td>
<td>10</td>
</tr>
<tr>
<td>2.4 Efficacy of different forms of subtitles</td>
<td>12</td>
</tr>
<tr>
<td>2.5 Statement of purpose</td>
<td>16</td>
</tr>
<tr>
<td>3. Method</td>
<td>17</td>
</tr>
<tr>
<td>3.1 Participants</td>
<td>18</td>
</tr>
<tr>
<td>3.2 Materials</td>
<td>21</td>
</tr>
<tr>
<td>3.2.1 Gathering background information and L2 proficiency data</td>
<td>21</td>
</tr>
<tr>
<td>3.2.2 Film</td>
<td>21</td>
</tr>
<tr>
<td>3.2.3 Vocabulary test</td>
<td>23</td>
</tr>
<tr>
<td>3.3 Procedures</td>
<td>25</td>
</tr>
<tr>
<td>3.4 Data processing &amp; analyses</td>
<td>27</td>
</tr>
<tr>
<td>3.4.1 Vocabulary test</td>
<td>28</td>
</tr>
<tr>
<td>3.4.2 Misspellings of target words</td>
<td>28</td>
</tr>
<tr>
<td>3.4.3 Eye tracking data</td>
<td>29</td>
</tr>
<tr>
<td>4. Results</td>
<td>30</td>
</tr>
<tr>
<td>4.1 Learning effect</td>
<td>30</td>
</tr>
<tr>
<td>4.2 Misspellings of target words</td>
<td>32</td>
</tr>
<tr>
<td>4.3 Eye tracking</td>
<td>34</td>
</tr>
<tr>
<td>4.3.1 Target words &amp; learning effect</td>
<td>35</td>
</tr>
<tr>
<td>5. Discussion</td>
<td>39</td>
</tr>
<tr>
<td>5.1 Learning effect</td>
<td>39</td>
</tr>
<tr>
<td>5.2 Misspellings of target words</td>
<td>41</td>
</tr>
</tbody>
</table>
5.3 Eye tracking 42
5.4 Suggestions for further research 44

6. Conclusion 46

Bibliography 48

Appendix A Questionnaire 52

Appendix B Adjusted subtitles 54
  B.1 Standard 54
  B.2 Reversed 59
  B.3 Intralingual 65

Appendix C Pre- and post-test 71

Appendix D Answering sheet 73
The Efficacy of Subtitles
Four Different Subtitling Conditions to Enhance FL Vocabulary Knowledge

Abstract. The present study attempts to answer the question whether there is a subtitling condition that is more effective for FL vocabulary acquisition than another. This is a relevant question for the language classroom, since using film as a medium to provide authentic input has gained popularity over the past few years. Several studies state that processing more than one modality (i.e. soundtrack and subtitles) increases cognitive load and works counterproductively. However, most studies point towards the efficacy of subtitling, and its beneficial qualities in FL vocabulary acquisition. Therefore, an experiment was set up to test the efficacy of three different subtitling conditions: standard, intralingual, and reversed subtitling. As a control condition, a non-subtitled condition was added. A film fragment was created that had one of the subtitling conditions. Participants were native speakers of Dutch and L2 speakers of English with an advanced level of proficiency. Participants were exposed to one of the subtitling conditions. Both before and after watching the fragment, they had to do a receptive vocabulary test, consisting of definitions of target words, to which participants needed to find the correct response. Even though a learning effect took place in each of the conditions, there was no significant difference between the four conditions. Small effects were discovered with the naked eye but statistics do not back this up. Finally, it is suggested that more, large-scale research is done on the subject of efficacy of subtitling and the difference between the three subtitling conditions.

Key words: Subtitling, dubbing, eye tracker, FL vocabulary acquisition, SLA, learning effect, reversed, intralingual, interlingual, standard, language education.

1. Introduction
The film industry is thriving: no expenses are spared to make the next award-winning motion picture. Millions are spent in the process of producing and many more of them are made during the time the films are shown in cinemas, on film festivals, and, eventually, when they are made available on DVD. There are many prizes to be won and many directors who want to attain one of them (i.e. an Academy Award, also known as an Oscar Award, or several of them). This does not only apply to directors, but also to actors, writers of the screenplay, the music writers, the editors, etcetera.
It is in human nature to want more, to want better, to want to be greater than anybody else. The sky is the limit. This sense of competition is what makes that excellent films have been produced in which actors, directors, and all people involved in the production all work together to create a piece of art. This is what leads to better results and even more realistic productions. Acting has become increasingly authentic over the past years, as have the so called ‘special effects’ which, in turn, lead to a better film experience. It can be said with certainty that the acting in modern films – with the occasional exception – has come to imitate naturalistic behaviour fairly closely.

This very fact could well account for the phenomenon that the medium has gained in popularity among the audience. Not only quality plays a part in this increasing popularity, but also the enormous diversity in genre, language, and culture. It would not be for a lack of choice that an individual does not enjoy watching films. Films come in all sorts and sizes: there is comedy, romance, thriller, drama, science fiction, history, war, short films, extremely long ones, etcetera. There are people who watch films and/or series for a hobby and spend several hours a week watching them and share their thoughts and ideas with others. Not only can films function as a topic of conversation, they also engage the viewer in the story, the language that is used, and scene setting. Even though the storyline may be fictional, and the language spoken may be made up (such as the Elvish languages in *The Lord of the Rings* (Kreeft, 2005)), a film can be educational. To turn back to the example of *The Lord of the Rings* (Jackson, 2001), there are people who have actually acquired some Elvish phrases and expressions and use these to communicate with other speakers of the language. An even more extreme and more telling example of language acquisition that happened because of exposure through film, is the Vulcan language used in *Star Trek* which is used during wedding ceremonies and similar situations. Admittedly, this is an example of taking a hobby to the extreme, but what the examples do show is that engagement and excitement can lead to acquisition of a new language, a notion that is supported by various studies that investigated the effects of films on language acquisition (e.g. Taylor, 2005; Bird & Williams, 2002; d’Ydewalle & van de Poel, 1999).

Due to the fact that film provides for an excellent source of authentic language input, and given that authentic input is one of the prerequisites for a learner to acquire a language (Krashen, 1985), film is a medium that is increasingly popular in the language classrooms (Bisson, van Heuven, Conklin, & Tunney, 2011).

Another important reason why film has taken on such an important role in the modern language classroom, is that films are engaging. According to Cszéér and Dörnyei (2005), this is considered to be one of the most salient factors, if not the most important one, in language acquisition. Together with authenticity of the language that students are
exposed to, engagement has made that film has become a popular tool in language education.

Films are not automatically a facilitating component in the process of language acquisition, however. For language classroom purposes, the various forms should be considered: there are films with or without subtitles (where written words translate what is being said), dubbed films (where voices translate what is being said), and films with subtitles in the same language as the spoken word. The choice that a teacher has to make for his or her students is an important one since it may be the case that not all types of subtitles – or the lack thereof – are suitable for that particular group of pupils, or that one type is more effective than another. Without knowledge about the effects the different types of subtitled films have on students, the use of films in the language classroom may fail to achieve its objectives. The present study will therefore attempt to discover what type of subtitling is best to acquire a language. Special attention will be paid to vocabulary acquisition since this is an aspect that can be grasped within the scope of this paper: “words, particularly nouns, are indeed the easiest building blocks in acquiring a new language; accordingly, it is not too surprising to find the first signs of language acquisition in the vocabulary” (d'Ydewalle & van de Poel, 1999, p. 240). The question of whether a specific type of subtitling has the ability to facilitate vocabulary acquisition was attempted to be answered by means of an experiment in which four different groups of people were exposed to four different versions of the same film (i.e. films that are accompanied by different subtitling conditions). Both before and after having watched the fragment, they were tested on their vocabulary knowledge in order to see how much their knowledge of target vocabulary had improved after having been exposed to the vocabulary in different ways.

Aside from the main research question mentioned above, two more questions will be addressed. The three questions that will be addressed are:

1. Is there a subtitling condition that proves to be the most effective for vocabulary acquisition?
2. Does the subtitling condition that students are exposed to affect the number of misspellings made in the target words?
3. Does spending more time looking at the subtitled area lead to a greater learning effect with regard to the target words?

In subsequent chapters, an overview of studies that have already been conducted will be given, after which the blanks in the current body of literature will be pointed out. These blanks will be attempted to be filled with the present experiment, of which the methodology
is provided in chapter 3. Afterwards, the results from this experiment will be presented and discussed.

2. Background
A great many studies have already been conducted on the issue of subtitling and its usefulness in the process of second language acquisition (SLA). Especially now that film has become an increasingly popular medium for teachers to use for language teaching, the effectiveness of using video in this manner requires examination. Below, an overview will be given of different existing views on the issue of subtitling and whether it should be used or not. At the end of the chapter, a link will be made between existing research findings and the study at hand which is an attempt to contribute to the discussion of whether and how to use subtitles in the language classroom.

2.1 Film in the classroom
It is argued by many that solely being exposed to explicit instruction (i.e. explaining rules and having students apply them to their language production (Ellis, 1994, p. 642)) is not enough to learn a second language (e.g. Ponniah, 2011; Krashen, 1985, 1989), but that a combination of comprehensible input and explicit instruction is the best way to teach a language. Given this information, it can be assumed that indeed some type of comprehensible input needs to be present in the language classroom. Instead of receiving only explicit instruction and in the process only learning dictionary definitions of words and not being able to actually use the words (Ponniah, 2011), it is important to receive context so that language becomes animate, instead of static information. In the end, what counts is knowing how to use a language, not knowing how to flawlessly recite dictionary definitions. Furthermore, Nation and Waring (1997) state that context is one of the most important factors in vocabulary learning strategies, provided that the vocabulary occurs on a frequent basis in the materials (Elley, 1989). Context is not acquired by dictionary definitions and calls for different types of materials, such as films, newspapers, books, etc.

Film, as mentioned before, is a medium that is used increasingly often in the language classroom to expose students to the target language (TL). Many teachers would agree that attention to a film is many times greater than attention to a teacher attempting to explain word order in questions, how to properly distinguish between the present perfect continuous and the present perfect tense, or trying to explain the meaning of a word by providing a dictionary definition. It has long been argued that solely providing explicit language instruction is not enough to learn a second language (L2). On the contrary, learners need to be exposed to the TL as much as possible and must, if they can, formulate
their own rules from the input they have been exposed to. This idea is supported by Krashen’s Input Hypothesis (Krashen, 1985; 1989), which implies that language learners acquire a language by reading or listening to comprehensible input. It is thought that learners do not need explicit instruction because they can themselves extract important information from the input and do this with the help of context, knowledge about the world, and knowledge of linguistic structures from other languages (Krashen, 1985, p. 80). The Input Hypothesis is likely to be applicable to infants learning to speak their mother tongue, a process during which explicit instruction is usually not received. Solely relying on comprehensible input during the process of acquiring a second language (L2), on the other hand, may make the road to (near)nativeness unnecessarily long. Krashen (1989) states that people acquire vocabulary through reading, and occasionally and subconsciously acquire several words. This statement is argued by other researchers, who agree with the notion that incidental acquisition of vocabulary takes place, but that this process is sped up when students’ attention is drawn to new phrases and vocabulary. The issue of ‘noticing’ plays the main part in this argument. It is, in the years following Krashen’s Input Hypothesis, argued by many that Krashen’s hypothesis has flaws. It is suggested that a combination of comprehensible input and explicit instruction is the ideal manner to quickly acquire a language (e.g. Schmidt, 1993; Rutherford, 1987; and Smith, 1991). Both Schmidt (1993) and Smith (1991) provide evidence that input needs to be noticed for it to become intake and to contribute to the learning process: only intake can turn into learning; input cannot. Therefore, if input is not noticed, it cannot become intake and, in turn, no learning can possibly take place (Schmidt, 1993, p. 218; Schmidt, 2010; Smith, 1991).

Aside from the necessity of comprehensible input, there is also a need for authentic input. Authenticity of input, Gilmore (2007) writes, is “inherently more interesting than contrived [input] because of [its] intent to communicate a message rather than highlight target language” (pp. 106-7). The fact that films communicate a message, or rather a narrative, is even more attractive to students because it engages them in the input. Engagement leads to motivation, which is, in turn, one of the most important individual factors to learn a language: without motivation, chances of succeeding in learning a language are slim (Csizér & Dörnyei, 2005). Engagement is a factor that is mostly absent in situations where a teacher explains grammar and teaching grammar may therefore be considered less effective when standing on its own. Furthermore, using a film in the classroom as either the centre of the lesson series or as a supplement to the regular curriculum leads to a focus on form: focus on the message rather than on the grammar that is used to convey that message (Gilmore, 2007, p. 107). When focusing on form (meaning) instead of forms (grammatical structures), authenticity of input, communication, and “the
occasional and overt study of problematic L2 grammatical forms” (Poole, 2005, p. 47) are covered, once again combining input with explicit instruction, which had already been proven to be the best form of language instruction (Norris, & Ortega, 2000; Lightbown, 2000).

All in all, the fact that films are being used in the language classroom more and more is a development that should be appraised. The question that remains, however, is how to do this in an efficient manner.

2.2 Overcoming language boundaries

Now that it has become clear that comprehensible and authentic input is a significant aspect in SLA and that using film as a medium providing this type of input is a good step towards good language instruction, it needs to be ascertained how film is to be made comprehensible for students with a native language (L1) different from the language used in the film. After all, when input is not comprehensible, there is no benefit in using the materials for instruction.

Essentially, the two most prevalent ways of overcoming the language boundaries between the students’ first language and the target language are dubbing and subtitling (Koolstra, Peeters, & Spinhof, 2002). Dubbing is a costly undertaking and the technique is therefore mostly used in countries that have a great number of inhabitants. Not only large countries like Germany and Russia opt for dubbing, but also countries that use censorship on a daily basis in media and politics. In dubbed film, the original language is usually no longer to be heard and the voice-over can therefore not be compared to what was originally being said (Koolstra et al., 2002). Even though this can be seen as an advantage when wanting do adapt or even change the message that is being conveyed by a film, in terms of language acquisition, showing dubbed films in the language classroom does not stimulate SLA. As can be seen from table 1 below, there are both positive and negative sides to dubbing; however, it will not be further discussed in the current study as it is excluded as a means to improve knowledge of a second language.

Table 1

Pros and cons for dubbing and subtitling (adapted from Koolstra et al., 2002, p. 344).

<table>
<thead>
<tr>
<th>(Learning) effects</th>
<th>Pro (+)</th>
<th>con (-)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stimulates reading development</td>
<td>Subtitling</td>
<td>Dubbing</td>
</tr>
<tr>
<td>Stimulates vocabulary acquisition in own language</td>
<td>+</td>
<td>+</td>
</tr>
</tbody>
</table>
Stimulates foreign language acquisition +
Unnaturalness through asynchronicity -
Bad translations because of condensation -
Incomplete or stretched translations -
Asks for high mental effort* -
Authenticity through hearing original actors +

* to be discussed in further detail in section 2.3

Subtitling, however, does provide the opportunity for SLA, an assumption that will be discussed in further detail in section 2.4. By providing subtitles (adding translations of spoken text at the bottom of the screen), a link between the target language and the mother tongue can be made. However, subtitles do not necessarily need to build a bridge between first and foreign language (FL): it is also possible to provide a transcription of the spoken text at the bottom of the screen.

2.3 Multimodal processing
Having to read subtitles that go with a video asks more attention from a viewer, or so many people think. In 1862, Wundt already mentioned there are processing costs when watching a subtitled film because of the simultaneous exposure to different kinds of input. This idea is supported by Broadbent (1958), who writes that only one sense (either sight or hearing in the case of film and subtitles) at a time can be put to full use. When the viewer’s attention is to be divided over two or more senses, the capacity to understand the input is compromised. It was believed that a learner should focus solely on reading or listening. However, as time went by, the idea that the human brain is capable only of using one sense at a time, has faded (d’Ydewalle, 2002). Instead, researchers like Navon and Grother (1979) have given rise to the idea that parallel processing is possible, i.e. that several senses can be used simultaneously to reach a better understanding of input. This is also known as multiple-resource theory (Basil, 1994), as opposed to the unitary-resource theory described by Kahneman (1973). This idea of a multiple-resource theory is also supported by d’Ydewalle (2002), who writes: “Partly due to the recent theoretical advances, almost everyone will no longer question the flexibility of the human mind in performing multiple tasks” (156).

However, no consensus has yet been reached on the issue. Quite recently, Taylor (2005) tested two groups of English speaking student of L2 learners of Spanish, first year students in one group, and third year students in the other group, to see whether there was a difference in level of comprehension of the film fragment that was showed in both groups.
This film had a Spanish soundtrack as well as Spanish subtitles. The study showed that especially the first year students experienced the subtitles as distracting and that it was too distracting to process the soundtrack, image, and the subtitles all at the same time. However, third year students had a much more positive attitude toward the subtitles. The study concludes by stating that “it seems that students with little background in reading and listening in the target language have difficulty attending to the three channels [subtitles, soundtrack, and image] and are confused or distracted by the use of captions” (Taylor, 2005, p. 426). Furthermore, it is a known fact that English speaking countries do not use subtitles as much as non-English speaking countries: most of the video materials available are already in English. L1 speakers of English are therefore at a disadvantage when being asked to watch a video with subtitles; as opposed to speakers of Dutch, to whom reading subtitles is perfectly natural. The difficulty in processing multimodal input may be explained by Sweller’s (1988; 2004) cognitive load theory (CL), a theory implying that there are several parts in the working memory and that each part has a limited capacity. Mayer and Moreno (1998) state, in relation to CL, that “meaningful learning [only] occurs when a learner retains relevant information in each store, organises the information in each store into a coherent representation, and makes connections between corresponding representations in each store” (p. 312). Apparently, this is difficult for beginning learners (Taylor, 2005). However, it should still be noted that 11% of the third year students in the experiment still found subtitles distracting, indicating that having to read subtitles is not experienced in the same way by everyone. In Taylor’s (2005) study, the third year students can be considered advanced learners of Spanish and the finding that several of them find subtitles distracting, may have implications for the question whether or not it is possible to process both spoken language and written language at the same time. Zhang (2013) supports CL and explains, with the help of theoretical background, why subtitles are distracting in the process of language acquisition: providing subtitles when learners should in fact be improving their listening skills, increases the cognitive load and will in turn lead to a decrease of intake from the auditory input (Zhang, 2013, p. 137-8). This study, however, merely uses already existing ideas to formulate a statement, and is not supported by any empirical evidence.

A recent study that is accompanied by strong empirical evidence, however, points out that subtitles are not distracting to students, but rather take away some of the frustration that occurs when listening to speech in a foreign language. Kruger, Hefer, and Matthew (2013) tested this empirically by having several students listen to a lecture with or without subtitles. They found, with the help of eye tracking, that there is a “significantly higher level of frustration for the unsubtitled condition” (p. 62). In the study, participants’ pupil dilations
were registered by an eye tracker, which provided information about the amount of frustration or stress the participant was under while watching the film fragment. Furthermore, levels of frustration were supported by the EEG (electroencephalography) data and it was concluded that “subtitles at the very least do not result in cognitive overload” (p. 62): frustration levels were higher for the unsubtitle condition, indicating that the initial idea that subtitles increase CL does not hold for this experiment, but rather that CL is increased when subtitles are absent. To enforce this idea, a study by d’Ydewalle and Gielen (1992) showed that switching between modalities (i.e. soundtrack and subtitles) does not take any effort and can even be considered automatic. This idea is supported by the automatic reading behaviour theory which implies that when something readable is in the viewer’s eyesight, it is being read automatically (d’Ydewalle et al., 1991; Cerf, Frady, & Koch, 2009). Cerf et al. (2009) provide evidence for the fact that viewers first fixate on faces, then on text, and then on the rest of the image.

Given the convincing evidence that points towards the notion that no real difficulty occurs while watching subtitled videos, it will, for the time being, be assumed that indeed adding subtitles to videos will not have a negative effect on a learner’s language acquisition. In fact, especially given the fact that many studies have confirmed that vocabulary is better learned when having access to subtitles (to be discussed in greater detail in the following section), it is highly unlikely that supporting subtitles are actually counterproductive in SLA. In fact, the current study provides evidence that there is indeed no negative effect on participants’ performance after having been exposed to subtitles, as opposed to participants who have not read subtitles (see chapter 5).

2.4 Efficacy of different forms of subtitles
Since previously mentioned studies have made it clear that processing subtitles is not too straining on the cognitive language learners’ abilities, the question that arises is whether subtitles aid SLA. The question of whether subtitles facilitate language learning has long been the focus of many studies. While watching a subtitled movie, the viewer is exposed to several forms of input: the visual image, the actors’ voices (soundtrack), and the translation or transcription of those voices (subtitles). In the previous section it has already been established that people are capable of using several senses to successfully process input. However, being capable of processing both visual and auditory input simultaneously does not automatically mean that it is also effective for learning a language: it merely describes human capabilities.

It is therefore important to investigate if and to what extent subtitles facilitate language learning. A subject also in need of discussion is the different types of subtitling
available and, of course, their effectiveness in SLA. For this cause, the three different types of subtitling will be discussed, after which an overview of different views on their efficacy will be provided.

The first, and most commonly known, type of subtitling is interlingual subtitling (standard subtitling). This is one of the easiest to produce and, when provided, most easily accessible type of subtitling to the ordinary audience. Everyone who can read in his or her native language, is capable of watching a film that is produced in any spoken language, provided that subtitles in the L1 are present. When this type of subtitling is used, the original soundtrack remains intact, while at the bottom of the screen, a (somewhat condensed) translation in the native language is shown.

The second, and less widely used, type of subtitling is intralingual subtitling. Here, subtitles do not contain a translation of the spoken word, but rather a (condensed) version, or transcription of what is being said. For language learners who have no knowledge of the language they are being exposed to while watching an intralingually subtitled film, this may be difficult. There is, after all, no available link to the first language. This particular link is used especially by beginning learners of a language (Weinrich, 1953). As learners reach advanced stages of proficiency, the L1-L2 link gradually fades away and knowledge of the target language is increasingly often used to interpret what is being expressed. This idea is supported by Lambert and Holobow (1984), who feel that beginning learners are not capable of accessing a FL without a link to the native language. In the intralingual subtitling condition, learners are exposed to both auditory and visual TL input, which would most probably lead to improvement in spelling as well as in pronunciation.

The final, and most foreign, subtitling type known today is reversed subtitling. In this condition, the original soundtrack is replaced by a soundtrack in the audience’s native language, while the subtitles translate this back into the target language. Learners are therefore not exposed to auditory input of the target language, but only to the written form. Table 2 (below) provides an overview of the three different types of subtitling and their expected effects on language acquisition.

<table>
<thead>
<tr>
<th>Condition</th>
<th>Languages used</th>
<th>Competences developed</th>
</tr>
</thead>
<tbody>
<tr>
<td>Standard</td>
<td>FL L1</td>
<td>L2 Pronunciation x L2 Spelling</td>
</tr>
<tr>
<td>Intralingual</td>
<td>FL FL</td>
<td>x x</td>
</tr>
</tbody>
</table>

Table 2

*Characteristics of the different types of subtitles plus control condition.*
Now that the various forms of subtitling have been discussed, it is time to look at studies that have researched their efficacy in language learning. It is widely accepted – with the exception of a few who believe that dividing attention over different modes of input is a strain on the brain – that subtitles have a positive effect on language learners.

Whereas many studies have only investigated the difference between bimodal (i.e. soundtrack and (standard) subtitles) and monomodal (only soundtrack) processing (among others, Hsu, 2014; Bird & Williams, 2002), only a few studies have been conducted that test the difference between two or three subtitling conditions. One of these is d’Ydewalle and van de Poel’s (1999) study in which the effects of two subtitling conditions on children was tested. In their study they have used Dutch audio with French subtitles and the other way around, and Dutch audio with Danish subtitles and the other way around, i.e. standard and reversed subtitles. They discovered that acquisition only really occurs when a standard form of subtitling is used, and that in the reversed condition, acquisition only occurs in the written form, since there is no auditory input when being exposed to a film with reversed subtitles. What this study most interestingly showed, however, is that adults appear to acquire more vocabulary when the native language is used in the soundtrack instead of in the subtitles (d’Ydewalle & van de Poel, 1999, p. 241). This indicates that adults gain more from being exposed to film in the reversed subtitling condition than by any other. This idea is supported by Danan (1992) and d’Ydewalle (2002). Whereas d’Ydewalle’s (2002) study focuses more on grammar acquisition and on only two out of three subtitling conditions, Danan (1992) tested which of the three subtitling conditions mentioned in table 2 above worked best for FL vocabulary acquisition. It was learned that the effectiveness of the reversed subtitling condition was significantly more effective than the other two conditions. In the study, L1 English speaking students of a second-year college course in French participated and the study showed that students in the reversed subtitling condition had twice as high test-scores than in the standard or intralingual condition. This finding is explained as follows (Danan, 1992, p. 522):

[R]eversed subtitling may be particularly beneficial because students have more time to process the foreign discourse and also benefit from the contextual knowledge gained from listening to their native language. In addition, once translation has linked the two verbal systems, students have
established more paths for retrieval and can benefit from visual traces as well as from two distinct sets of verbal traces.

The last part of the quotation is explained by the bilingual dual coding theory (Paivio, 1986), which indicates that when someone has two paths to retrieve a word, retrieval is done more effectively, and more quickly, than if only one path were to be established. It is furthermore suggested by Danan (1992) that it may not be as effective as many people think to expose learners to an FL without any links to their native language, unless these learners are exposed to the FL for longer periods of time in their daily lives. This goes against Weinrich’s (1953) and Lambert and Holobow’s (1984) view that only more advanced students have the ability to access an FL without interference of their L1. In Danan’s experiment, participants were of both beginning and advanced proficiency levels, indicating that his findings also apply to the more advanced ones. Participants in Danan’s experiment were told beforehand that they were to be tested on the vocabulary presented in either the subtitles or in the soundtrack. On top of that, they were asked to indicate which words they had known before watching the film fragment so that a comparison could be made between pre- and post-test. This, however, is a very subjective manner to do this type of research: an actual pre-test would have been a more objective measure to analyse any noticeable differences between prior and acquired knowledge.

An eye tracking study by d’Ydewalle and de Bruycker (2007) shows that much more time is spent in the subtitled area when these were in the native language (i.e. standard subtitling) and that subtitles in reversed subtitling condition were more often, but not altogether, skipped. It was concluded that there is a general preference for native subtitling. However, these results were based on eye tracking data only, and therefore does not say anything on the effectiveness. If, however, a pre- and post-test would have been administered, the study could have provided insight into the efficacy of subtitles. Solely relying on eye-movement data is not enough to draw conclusions on the effectiveness of a subtitling condition.

Bisson, van Heuven, Conklin, and Tunney (2012) conducted an extensive study on all three kinds of subtitles with the help of eye tracking and, as opposed to d’Ydewalle and de Bruycker (2007), did incorporate a vocabulary test to add meaning to the experiment. The initial question that was posed in this study was to what extent the different types of subtitles are processed. The same conclusion was drawn in Bisson et al. (2012) as in the d’Ydewalle and de Bruycker’s (2007) study: participants spent considerably more time reading the subtitles when the soundtrack is in an FL. It was mentioned, however, that regardless of the fact whether subtitles were provided in FL or L1, subtitles are always read
and that subtitles were never there without cause. The fact that subtitles are read regardless of the fact that viewers understand the spoken message, was explained by a link between salience of subtitles and automatic reading behaviour (Laberge & Samuels, 1974; Samuels, 1994; d’Ydewalle et al., 1991): whenever words appear on the screen, the viewer’s attention is drawn towards them and they are read automatically. The vocabulary test in Bisson et al.’s (2012) study – an auditory word recognition test – was done by all participants, while the film was not watched by everyone. Participants were split up into two groups, one that did watch the film, and one that did not. The group that was exposed to the film participated in the eye tracking experiment and had to do the vocabulary test afterwards, while the other group was not exposed to the film fragment and only had to do the vocabulary test once. Participants were native speakers of English had no prior knowledge of the FL used in the experiment: Dutch. The vocabulary test that participants had to do after having watched the film showed vocabulary acquisition, but since this was also the case for participants who had not watched the film, this effect was not established by any type of subtitling, but could well have been an effect of chance. It is observed that a more sensitive measure needs to be incorporated in further studies.

Seeing as there are many different views on subtitles and their effectiveness in FL vocabulary acquisition, and that different studies have either done research in which the method is questionable, or have gained results that are not applicable to the research question at hand, more research is required.

2.5 Statement of purpose
The aforementioned studies have put forward inconclusive results with regard to the efficacy of subtitles (are people capable of processing multimodal input? are subtitled videos an effective medium in SLA?) and the efficacy of different types of subtitling. The research in these studies mainly focused on the question whether subtitled videos are more facilitative to vocabulary acquisition than videos without subtitles. Given the fact that there are three different types of subtitles instead of just ‘a video with subtitles’, the question is not whether subtitles are facilitative to SLA, but which type of subtitling is the most helpful in the process.

Even though there have been studies that included all three subtitling conditions (e.g. Bisson et al., 2012), these studies have not been successful in producing conclusive answers to the question which condition is best for vocabulary acquisition. Since there are many studies that point toward a difference between the efficacy of the different subtitling conditions, the current study will make another attempt at providing an answer to this question.
Given that previously conducted studies including all subtitling conditions (instead of solely one or two) lean towards the idea that reversed subtitling is the most effective in SLA, the hypothesis that will be tested in this study is that reversed subtitling results in the best learning effects for vocabulary acquisition of the target language. Since the effectiveness of subtitling has been proven by each study – with the exception of only a few – it will be expected that the group that is exposed to only target language audio, shows a smaller increase in vocabulary knowledge. If the opposite result is found, then the monomodal input is easier to process than the multimodal input, supporting the CL theory (Sweller, 1988; 2004). Furthermore, since in the present study only advanced learners of English are tested, the intralingual condition may prove to be the most effective subtitling condition because advanced learners make little use of the links between L1 and FL (Weinrich, 1953).

The second research question that will be discussed is if subtitling condition affects the number of misspellings made in the target vocabulary. The second hypothesis that will be tested, is that participants exposed to Dutch subtitles or no subtitles instead of English ones, will show a greater number of misspellings in their post-test when a they use the newly acquired target word. This is expected since participants in this condition will not have seen how the word is written before and have to write the new word solely on the basis of auditory input (Danan, 1992).

The third and last question that will be addressed is whether there is a relation between the time spent looking at the subtitled area and the learning effect pertaining to the target words. Since the theory exists that the more often a learner comes across a word in context, the more likely it is that he or she acquires the word (Krashen, 1985; 1989; Nation, & Waring, 1997), it is hypothesised that the amount of time spent in the subtitled area has a positive effect on the learning effect.

3. Method
To investigate the research questions and hypotheses mentioned in section 2.2 above, an experiment was set up that requires participants to watch a video fragment accompanied by one of the aforementioned subtitling conditions (i.e. standard, intralingual, or reversed see also table 2 above) or a control condition which consists of a video fragment with FL audio but no accompanying subtitles. An experiment was set up that provided the chance to take a look at the learning effect as a result of the subtitling conditions. It is expected that some conditions have a greater effect on vocabulary learning than others and it can therefore be assumed that the extent to which participants improved their vocabulary knowledge is dependent upon the subtitling condition they were exposed to (the independent variable).
In the following sections, the method is described in greater detail. First, in section 3.1 an overview of the participants will be given, after which the materials that were used will be discussed in section 3.2, followed by a description of how these materials were used in section 3.3 and by how the data were analysed in section 3.4. Lastly, the results from the experiment will be presented and discussed in chapters 4 and 5.

### 3.1 Participants

For the pilot study, a dry-run before the actual experiment would take place, there were five university students who volunteered to participate. Students were not selected with any background in mind: the data would not be analysed so results were not important.

For the actual experiment, a total of twenty students (sixteen of whom are female) with different study subjects and varying numbers of years in higher education behind them, participated. In order to work with a participant group of more or less the same cognitive abilities, the choice was made to only ask university students to participate. In figure 1 below, an overview of the different study subjects is visualised, showing that the lion’s share of the participants have studied – or are still studying – English or another language.

![Figure 1: Study subjects among the twenty participants.](image)

All participants have a study topic that falls within the confines of the alpha subject area, so that potential differences between participants cannot be due to the contrast between alpha and beta students (e.g. beta students being expected to have less language aptitude than alpha students).
Selection of participants was done in two ways. On the one hand, there were nine students who had to partake in several linguistic experiments during their studies and had to write reports on them. All of these students have English as their major subject. The remaining eleven participants were selected through several other methods: student advisors from different departments in the university were contacted who, in turn, sent an invitation to their students. Others read a message on Facebook or heard from other students who had already partaken in the experiment that participants were still needed. This group of students therefore consists of students with various backgrounds. All students were aware that there would be no financial compensation for participating in the experiment, even though it would take approximately 1.5 hours. It was mentioned, however, that three € 5,- gift cards would be arbitrarily distributed among participants. The absence of a full compensation for the time spent in the experiment means that students who did partake were motivated to do so because of the study and not because of the money.

Since the only three required characteristics participants had to have were 1) having Dutch as a first language and having knowledge of English, 2) having a university level of education, and 3) having a topic of studies that falls within the confines of alpha subjects, any other differences between participants are vast. For example: several students only know three languages (Dutch, English, and German), whereas others speak Dutch, English, German, French, Spanish, Frisian, and Swedish. In figure 2 below, an overview is given of all the languages spoken by the participants. In total, there are four participants who were raised bilingually, either from birth, or from a later age. In three of those cases Frisian was the other language. In the fourth case, this language was English, suggesting that this participant has an equally high proficiency in both Dutch and English, presumably (near) native in both languages.

Another variable that makes the group of participants diverse is age. The age ranges from 19 to 41, with a mean of 22. The lion’s share of the participants have an age between 19 and 22. Furthermore, only six out of the twenty participants have spent time living abroad in an English speaking country. This time ranged from ten weeks to three non-consecutive years. The rest of the students speak or read English at a regular basis – mostly for studies – but use Dutch for communication and reading for the greater part of the day.

The positive side of having a group with much variability between the test subject is that the sample is more representative, meaning that results from this experiment are more applicable to the real world, rather than just a select group of people.
Participants were distributed among the four different conditions relatively evenly, as can be seen from table 3 below.

<table>
<thead>
<tr>
<th>Subtitling condition</th>
<th>Number of participants</th>
</tr>
</thead>
<tbody>
<tr>
<td>Standard</td>
<td>5</td>
</tr>
<tr>
<td>Intralingual</td>
<td>5</td>
</tr>
<tr>
<td>Reversed</td>
<td>6</td>
</tr>
<tr>
<td>Control</td>
<td>4</td>
</tr>
</tbody>
</table>

The reason that the number of participants in the reversed condition is larger than in the rest of the conditions is because of the drop-out rate: it turned out that during the experiment the eye tracker was not fully capable of following the eye movements of one of the participants. It was therefore deemed necessary to have one more participant in the reversed condition and one less in the control condition because eye tracking data will be less valuable in the latter group since there are no subtitles available in that condition. Unfortunately, there were 5 other cases in which the eye tracker failed to flawlessly record eye movements.
3.2 Materials

3.2.1 Gathering background information and L2 proficiency data
Prior to the experiment, each participant was asked to complete a sociolinguistic background questionnaire (see Appendix A). The language that was chosen to formulate the questions is Dutch. This decision was made so as not to prime participants during the pre-stages of the experiment to the target language, English. The decision holds throughout the entire experiment, where the language used for instruction is Dutch. The questionnaire functions to provide insight into basic information such as age, education, and the knowledge of languages. The part in which participants are asked about the languages they speak and how well they master them is particularly interesting. Not only does it provide insight into how well languages are spoken, but also into participants’ motivation to learn languages. Students in secondary school in the Netherlands need to learn two foreign languages: both English and either French or German. When a participant mentions that he or she knows more than two foreign languages, this informs us of the willingness to learn new languages.

For the experiment itself, several test forms and other tools and devices were used. First, to establish the participants’ proficiency level, DIALANG (Dialang Partnership, 2006) was used. DIALANG is a computer testing programme that is used primarily as a diagnostic tool for students to know what language proficiency level they are at, according to the CEFR scales (Council of Europe, 2001, also mentioned by Chapelle, 2006). Even though the Dialang Project has not been updated since 2006, it can still be considered a rather reliable tool to decide upon students’ proficiency level. During the experiment, participants were required to take two tests in DIALANG, both for reading and for listening, since these are the two skills that are used while watching a subtitled film.

3.2.2 Film
The part of the experiment where the actual data was collected, is subdivided into three parts: pre-test, film fragment, and post-test. The film fragment and the eye-tracker that were used will be discussed first, after which the pre- and post-test will be elaborated on.

Selecting a suitable film for this experiment was the first issue to hurdle. The film would have to be engaging for approximately ten minutes, would have to appeal to both male and female adults, must contain several words that are probably not known by most participants, must contain unknown words that are used repeatedly, and watching the film must not be an annoying experience for participants exposed to the dubbed version (reversed condition) of the film (Bisson et al., 2012). Taking these five points into consideration, the film that was opted to use was The Pirates! Band of Misfits (Lord, Sproxton, & Lockhart, 2012). This is a relatively new animated film, and less well known
than the popular Pixar movies such as *Shrek* and *Finding Nemo*. The fact that a less popular film was chosen is because then the chance that the film is recognised by many participants is smaller. Furthermore, an animated film is available in different spoken languages, which mainly is a practical reason to opt for an animated movie. A factor that plays an important role in choosing for an animated film is that it does not matter much which language is used as the spoken language, inasmuch as lips move relatively synchronously to any language that is used for the soundtrack which, in turn, leads to less annoyance. This would not have been possible with a film that was acted by real-life actors since then the lips would most definitely have moved differently when the film would be dubbed in any language other than the original one. One important note on the version of the film that was dubbed (reversed condition) is that, even though the language used is Dutch, the accent is Flemish: Flemish and Dutch voice actors usually work together on dubbing assignments in the Netherlands.

After the film had been chosen and a fragment – a fragment that would be understandable without receiving any background information – of twelve minutes, was picked out, subtitles had to be made to fit the different subtitling conditions. While the original subtitles that went with the DVD were inaccurate and sometimes incomplete, all of them were rewritten and made in such a way that words that were spoken in the version with English audio, were also typed in the version with the reversed subtitling condition, so as to make sure that each test subject would be exposed to the same English target language. The programme that was used for this Subtitle Edit (Softonic, 2010). Into this programme, the prerequisites of subtitles were already programmed, so that the subtitles were conform the commonly accepted rules (adapted from d’Ydewalle, van Rensbergen, and Pollet (1987)).

- A subtitle cannot contain more than two lines;
- There cannot be less than three characters in a subtitle;
- A line of subtitles can hold no more than 36 characters (spaces included);
- Each subtitle must be displayed for at least 1.5 seconds;
- Between two subtitles, a blank of at least .25 seconds should be displayed;
- One-line subtitles cannot be displayed for more than 3.5 seconds;
- Two-line subtitles cannot be displayed for more than six seconds.

Even though there are many more standards to which subtitles must adhere, these are too elaborate for the scope of this project and therefore only the seven key points mentioned above have been included here. The final products of the subtitles can be found in Appendix B (B.1 standard, B.2 reversed, B.3 intralingual). A slightly different version of English
subtitles (Appendix B.2 and B.3) was used for the reversed and intralingual conditions. This has to do with the fact that the subtitles needed to be a near exact transcript from what was being said in the soundtrack in the intralingual condition. The wording is slightly different – except for the target vocabulary – and the timing is not exactly the same either. These are only minor differences, but are worth noting.

After four separate AVI files for the different subtitling conditions had been made with the dimensions of 1280×1024 pixels, all files were transferred onto the computer attached to the eye tracker. The programme that was used to record the eye movements and to process the files is Tobii Studio (Tobii Technology, 2013). In this programme, areas of interest (AOIs) were made for each film fragment separately. These AOIs were made both for the entire subtitle area as for the twenty target words (see section 3.2.3 below for more details on the target words). This would provide for the opportunity to analyse the data more easily since it would be visible from the start whether participants looked at the target words and if so, for how long and how often. It will then also be more easy to compare the effect of the different conditions with each other.

The eye tracker that was used to record the eye movements during the experiment is built in into a computer screen that is administered through a regular computer elsewhere in the room. Even though there are portable eye trackers available on the market, this type of technology was not available at the University of Groningen. Using a non-portable eye tracker means that participants have to hold still while watching the film as much as possible, since if they move too much and then return back to their original position, the eye tracker may have lost the signal and will no longer record the eye movements properly.

3.2.3 Vocabulary test

Solely gathering information on whether a participant has looked at a certain word and for how long he or she has done so, is not enough to answer the question which type of subtitling is best suited for vocabulary acquisition. Information is needed on whether the participant has learned a word after having seen it, or whether he or she has only seen it, but has not remembered it. Therefore, as a final part, a vocabulary test was created to determine participants’ knowledge of the target vocabulary knowledge both before and after watching the film fragment.

During the pilot study, it became clear that having a vocabulary test that consists solely of target words is not an option: test subjects mentioned that they figured out what the film was going to be about. Also, it showed that they became frustrated when they did not know the words they were supposed to write down. Furthermore, they mentioned that while watching the film, they paid specific attention to the target words they had
encountered in the vocabulary test. For these reasons, the set up of the vocabulary test was slightly changed after the pilot study.

A total of twenty target words were chosen from the film fragment. In order to try and make it less obvious that the words asked for come directly from the film fragment, another twenty words were added to the test as fillers so as to distract participants from the target words. An overview of words can be found in table 4 below.

**Table 4**

*List of target words and fillers.*

<table>
<thead>
<tr>
<th>Target words</th>
<th>Fillers</th>
</tr>
</thead>
<tbody>
<tr>
<td>Watery grave</td>
<td>Umbrella</td>
</tr>
<tr>
<td>Untold riches</td>
<td>Teacher</td>
</tr>
<tr>
<td>Trophy</td>
<td>Skirt</td>
</tr>
<tr>
<td>Squid</td>
<td>Rain</td>
</tr>
<tr>
<td>Shanties</td>
<td>Police</td>
</tr>
<tr>
<td>Lubbers</td>
<td>Pillow</td>
</tr>
<tr>
<td>Inclement</td>
<td>Pen</td>
</tr>
<tr>
<td>Hold</td>
<td>Newspaper</td>
</tr>
<tr>
<td>Hoist</td>
<td>Key</td>
</tr>
<tr>
<td>High seas</td>
<td>House</td>
</tr>
<tr>
<td>Hare-brained</td>
<td>Horse</td>
</tr>
<tr>
<td>Grant</td>
<td>Hair</td>
</tr>
<tr>
<td>Cutlass</td>
<td>Goal</td>
</tr>
<tr>
<td>Confound it</td>
<td>Giraffe</td>
</tr>
<tr>
<td>Brine-soaked</td>
<td>Flower</td>
</tr>
<tr>
<td>Booty</td>
<td>Dog</td>
</tr>
<tr>
<td>Barnacle</td>
<td>Computer</td>
</tr>
<tr>
<td>Baboon</td>
<td>Car</td>
</tr>
<tr>
<td>Avast</td>
<td>Book</td>
</tr>
<tr>
<td>Anecdote</td>
<td>Apple</td>
</tr>
</tbody>
</table>

For the pilot study, a vocabulary test was created that included both sentences in which a word was left out and descriptions of words that the test subjects had to write down. However, it was decided that it would be better to create a vocabulary test in which all test items have the same format. So, for the final experiment, a vocabulary test was created in
which each of the words in table 4 above was given a description. This was done with the help of the dictionary definition. The target words were put in reversed alphabetical order and their definitions were put in the vocabulary test (see Appendix C).

While more than one word may be considered a correct response to a definition, there is only one response that will receive full points (see section 3.4.1 for more information on scoring). In order to point the participant in the right direction, the first letter of each word was already provided on the answering sheet (see Appendix D).

Of course, the intent of this experiment is to see whether subjects learn new vocabulary. This cannot be concluded from only one test taken after having seen the film fragment, as was done by Danan (1992). The score from the post-test should be compared to a score from a test made prior to watching the movie in order to see whether the film has had an effect on the participants. For the sake of being able to judge whether test subjects have gained knowledge of vocabulary, the pre-test should test the same words as the post-test. Otherwise chances are that words in the pre-test are less well-known than words in the post-test and if a difference occurs between the two tests then, this will not be because of the film fragment. On top of using the same target words, the same test-format should be used so that differences in performance between pre- and post-test cannot be due to the fact that some students are better with one format than with another. The downside of this way of testing vocabulary knowledge, is that after having done the pre-test, participants will have been primed to listen for the words in the film fragment that they had minutes before encountered in the test. This, however, is a problem that cannot be avoided and needs to be taken for granted in this case.

3.3 Procedures
The experiment was carried out in two stages. Because it was unclear how much time was needed for the entire procedure and whether there were any unforeseen problems, it was opted to first do a pilot study ere the final experiment would be carried out. The procedures of the final experiment will be discussed in greater detail below.

For the first stage of the experiment, participants were one by one met in a quiet teacher’s office – in which some of the times the teacher was also working quietly – and were asked to sit down in front of a laptop. After participants were installed and were ready to start, an explanation was given of how the vocabulary test needed to be completed. They were explicitly asked to leave no open spaces and even when they did not know the word that they were required to write down, they were asked to fill in a word that made no sense at all. It was explained that it would be better to write a nonsense word than to write no word at all. When the instructions were clear, participants were told they had ten minutes to
complete the vocabulary test. This time was administered and after the ten minutes were over, the sheet with the test items was taken away while participants had to keep hold of the answering sheet a while longer because the results from the proficiency test that they had to take afterwards, needed to be written down on that same sheet.

The final part of the first stage consisted of a language proficiency test. Participants were instructed that they had to finish two tests within an hour and were told that that meant they did not have a lot of time to think of a correct answer. Furthermore, it was explained that it was not very important for the study that they did their utmost bests, but rather that the test would provide insight into their proficiency levels: to see whether all students were at approximately the same level so that results would be comparable. Participants were also explicitly instructed to write down their results from a test immediately after they had finished the test because once having clicked the screen with the results away, there was no going back and the test would have to be retaken. When this was all made clear and students understood what was asked of them, DIALANG was launched and instructions were switched to Dutch so as not to prime them for the target language. A quick explanation of the buttons that were required to be used during the test was given and then the test could be started.

Each student first had to complete the listening test, for which part headphones were available so that this part of the test would not disrupt the quiet in the room if another participant happened to arrive early, or another participant finished late, in which cases there were two students working at the same time. Several test subjects finished early in the reading part because they were self-admitted fast readers, while others took a little more time. After having finished the listening part and having written down the proficiency level on the answering sheet, the reading test was launched. Overall, the entire first stage of the experiment – the DIALANG tests including the vocabulary test – took no longer than 75 minutes.

For the final part of the study, watching the film fragment and completing the post-test, which was the same vocabulary test as the pre-test, participants had to go to another room in the building where the eye tracker was situated. In this room, the eyelab, the eye tracker was set on a table with a comfortable chair. Next to this table there are two cabinets that shield the computer on which the test materials were prepared for the test from the test-taker. Students were only to see the eye tracker and were not allowed to go past the cabinets. Instead, when they entered the room, they put their belongings on an empty chair and were asked to take place behind the screen with the built in eye tracker. Participants were asked to sit in a way they could keep still for about 12 minutes, while the programme Tobii Studio was started up so that the eye tracker could be adjusted to the test-taker. This
was done with the help of a tool in Tobii Studio that shows the position of the eyes and whether the signal that was received by the eye tracker was sufficiently good. Participants were required to sit at approximately 60 centimetres from the screen and their eyes should be approximately in the middle of the screen. When the eye tracker was properly adjusted to these conditions and headphones were put on, students were told to hold still while a five-point calibration session would start, during which students’ eyes had to follow a dot that moved along the screen. When the calibration was successful, students were asked to no longer move and sit still until the film fragment was over. If it was not successful, another calibration was done, until it was acceptable. An instruction was provided that told students that they had to watch the film as though they were watching a film at home, except for the fact that they could not move, or as little as possible. It was mentioned that they would have to answer a few questions afterwards, but it was not said that these questions would be the same ones that they had answered before. When all this was clear, the film fragment was started. This took approximately twelve minutes, after which the Tobii Studio project was stopped and the eye tracker ceased to record eye movements.

Immediately after having watched the film fragment, another vocabulary test and empty answering sheet were handed out. Again, participants received ten minutes to complete the test. When the time was up, several students asked what the experiment was about exactly and some asked to be shown what the eye tracker actually recorded. When an explanation was given, students were offered a candy bar, after which they were thanked to have participated in the experiment. The entire experiment lasted 75 to 100 minutes. In cases that it lasted 75 minutes, participants were really quick in doing the DIALANG tests. After all twenty participants had done the experiment, the three promised gift cards were sent through e-mail to three arbitrarily chosen participants.

Even though the experiment took quite a long time and there was no financial compensation other than a chance of receiving a gift card, there was only one participant who called in sick before the experiment. Nobody quit during the experiment and most of them found it a nice experience, especially to find out what their proficiency level actually is. Furthermore, they found it an interesting experience to participate in an experiment in which an eye tracker is used.

3.4 Data processing & analyses

The three research questions posed before (is there a difference in learning effect between the different subtitling conditions; does subtitling condition affect misspellings of target words; does spending more time in the subtitle area lead to greater learning effects?) will be
addressed in this section. An explanation will be given of how the data were processed and what the expected answers are.

3.4.1 Vocabulary test
The analysis of the data took place in two stages. First, the data from the pre- and post-tests were analysed in order to check for significant differences. In this study, the independent variable consists of four, and the dependent variable consists of two levels, so to see whether there is a significant difference between pre- and post-test in the different groups, a repeated measures test will be carried out in SPSS. A repeated measures test has the ability to analyse the difference between three or more groups and their test results from tests that were taken at different points in time. The hypothesis that will be tested is that the participants in the control condition will show the smallest learning effect, while the participants in the intralingual condition will show the greatest learning effect. The standard and reversed conditions are expected to have a similar learning effect (see table 2 above for the different subtitiling conditions).

The data from the pre- and post-test were scored according to the level of correctness. Both choosing the correct word and writing the word correctly was rewarded. Table 5 below displays the grading scales. The scores for each of the words were added up and this calculated value is the final score.

Table 5
Rating scales for the vocabulary tests.

<table>
<thead>
<tr>
<th>Grade</th>
<th>Degree of correctness</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>No answer or answer in a language other than the target language</td>
</tr>
<tr>
<td>1</td>
<td>Unrelated word and incorrect spelling</td>
</tr>
<tr>
<td>2</td>
<td>Unrelated word and correct spelling</td>
</tr>
<tr>
<td>3</td>
<td>Related word and incorrect spelling</td>
</tr>
<tr>
<td>4</td>
<td>Related word and correct spelling</td>
</tr>
<tr>
<td>5</td>
<td>Correct word and incorrect spelling</td>
</tr>
<tr>
<td>6</td>
<td>Correct word and correct spelling</td>
</tr>
</tbody>
</table>

3.4.2 Misspellings of target words
A second test was conducted to test whether the number of misspellings of target words differs significantly among the conditions. This will be done with a repeated measures analysis again, since misspellings of both pre- and post-test will be used. Data that will be
included in this part of the test are only those answers that have received 5 points according to the rating scale in table 5 above. Since in some cases it occurred that words that were not target word were spelled correctly in the pre-test and incorrectly in the post-test, including these results would provide a false picture of the effect of the different conditions. The hypothesis tested with this analysis is that most incorrect spellings of target words occur in cases where participants have not been exposed to the correct spelling (i.e. participants exposed to the control condition and the standard condition).

The \( \alpha \)-level for both repeated measures analyses is set at .05. This leaves an error-rate of 5\%, which is acceptable within the domain of applied linguistics.

3.4.3 Eye tracking data

The second stage of the analysis consists of the data that were collected by the eye tracker. To see whether participants have picked up words from the soundtrack or whether they have actually seen the words in the subtitle area, the data from the post-test and the data from the eye movements will be compared. For this stage of the analysis, the eye tracking data that provide information about the duration of fixations on different AOIs are used. The reason why fixations data and not visitations data are used is that with a fixation it is more plausible that test subjects have noticed the word because they have looked slightly longer in that particular AOI. This will be done to see whether this may have had an effect on the post-test scores. A Kruskal-Wallis analysis was carried out to analyse this. In addition to looking at the time spent in the subtitled areas, a closer look will be taken at three target words (trophy, watery grave, and inclement) to see how eye tracking data can account for differences in learning effect. An increase in score is a learning effect when the answer in the post-test is either the correct word spelled correctly or a misspelled version of the correct word. If an increase in score is visible, but the post-test answer is not the target word, or a misspelled version thereof, there is no learning effect since it was not caused by the audio or subtitles in the film fragment.

For this part of the analysis, the data from the eye tracker were used. Data that were used consist of times spent looking at certain AOIs, discussed in section 3.2.2 above. The data were extracted from Tobii Studio and put in an excel file so that the data could be processed from other computers.

---

1 Misspellings in words that do not occur in the film fragment are not relevant to the question that will be attempted to answer: the question whether there is a difference between conditions as far as correct spelling of target words is concerned.
4. Results

4.1 Learning effect

Homogeneity of variance was assumed for both the pre- and the post-test since Levene’s test showed non-significant results. After having inspected the histograms for each condition, it was also assumed that a roughly normal distribution was present. Since both assumptions have been met, the repeated measures ANOVA was carried out.

There was no significant overall difference between the scores in the four different conditions, \( F (3, 16) = .148, p = 0.929 \), but there was a main significant effect of time on score, \( F (2, 55) = 163.2, p < 0.001 \). Students scored significantly better on the post-test (\( M = 83.75; SE = 14.16 \)) than on pre-test (\( M = 67.10; SE = 17.20 \)). Figures 3 and 4 below show the difference between the four conditions and the improvement from pre- to post-test respectively.

![Scores on the pre- and post-test, displayed per condition.](image)

Figures 3 and 4 below show the difference between the four conditions and the improvement from pre- to post-test respectively.

As can be seen from figure 3, the increase in vocabulary test scores is approximately the same in each condition: the learning effect is approximately of the same size in all conditions. This idea is amplified by figure 4 below, in which the improvement is visualised.
A repeated measures ANOVA showed that there was no significant overall difference between the four groups, $F (3, 16) = 1.05; p = .399$. The number of blanks differed significantly between pre- and post-test: there is an interaction between number of blanks and time, $F (1, 16) = 7.213, p < 0.05$. Participants left significantly more open spaces in the pre-test ($M = 2.65; SE = 3.48$) than in the post-test ($M = 1.20; SE = 2.09$). One participant went from ten open spaces in the pre-test, to zero in the post-test, one went from ten in the pre-test, to five in the post-test\(^2\). A visualisation of the interaction between open spaces in the pre-test and open spaces in the post-test can be found in figure 5 below. From the figure, it can be seen that a particularly large difference can be found in the control condition. Even though no significant difference was found in the number of blanks between the different subtitling conditions, from figure 5 it can be deduced that the difference between the number of blank spaces in the pre-test is considerably higher than in the post-test. This particular difference is smaller than in the rest of the conditions.

\(^2\)All the while there was only one participant who showed a increase of one open space in the post-test.
4.2 Misspellings of target words

Levene’s test for homogeneity of variance showed non-significant results for homogeneity of variance and with this knowledge, one of the conditions for a repeated measures ANOVA is met. Having created histograms showing distributions that approximated normality, it is safe to do a repeated measures test for these data.

The number of misspellings did not differ significantly between the four conditions, $F(3, 16) = .626, p = 0.609$, however, there was a significant effect of time on number of misspellings, $F(1, 16) = 13.36, p = 0.001$. Students made significantly more errors in the post-test ($M = 1.4; SE = 1.27$) than in pre-test ($M = .35; SE = .59$). Figures 6 and 7 below show the difference between the four conditions and number of misspellings during the pre- and the post-test.
As can be seen from figure 6 above, there seems to be a large difference between the standard and reversed conditions; however, there is no significant difference between the conditions. To better visualise the increase in misspellings from pre-test to post-test, figure 6 below displays the differences between the two test. In accordance with the statistical results it can be deduced that the differences are indeed too small to be significant.
4.3 Eye tracking

Since the data from the time spent in the subtitled areas are not normally distributed, it is not possible to use a parametric statistical test to see whether there is a significant difference between the four conditions. Therefore, it was chosen to use a Kruskal-Wallis H analysis to find out whether participants spent more time in the subtitled area for one of the subtitling conditions. The test showed that there was no statistically significant difference in time spent in the subtitle area and the different subtitling conditions, $\chi^2(3) = 5.53$, $p = .137$. Even though the boxplot displayed in figure 8 seems to visualise a difference between the conditions, according to the Kruskal-Wallis H, this difference is not significant.

Figure 7: Visualisation of misspellings per condition per vocabulary test, and increase of misspellings displayed per condition.
The control condition was included as a control condition. After all, there are no subtitles in this condition. Still, an AOI was created in this condition, to find out whether participants looked at the subtitled area even if there were no subtitles. If the time spent looking at the subtitle AOI were the same for all conditions, then participants did not look at the subtitles per se, but at the pictures from the motion picture. Since this is not the case, participants did look at subtitles when they appeared on the screen.

4.3.1 Target words & learning effect
To see whether eye movement data can say anything about the learning effect of three target words, three target words will be looked at more closely. These are trophy, watery grave and inclement. It must be mentioned at this point that at times it occurred that eye movements were not tracked at all points in time while the film fragment was playing. This means that results displayed in this section may be inaccurate.

Starting with trophy, a word that is known in the pre-test by 18 participants in the pre-test, is known by 20 participants in the post-test. In the pre-test, however, there are 5 cases in which the target word has been misspelled (i.e. received 5 instead of 6 points). The word trophy occurs several times in the subtitles and is visible for a total of 8.583 (standard, ‘trofee’ Dutch translation of ‘trophy’), 7.022 (reversed), or 6.822 (intralingual) seconds, and is pronounced in the soundtrack on three occasions per film fragment.
Table 6

Test scores on trophy and time fixated on AOI. Time is in seconds.

<table>
<thead>
<tr>
<th>Participant</th>
<th>Score pre-test</th>
<th>Score post-test</th>
<th>Time spent in AOI</th>
<th>Condition</th>
</tr>
</thead>
<tbody>
<tr>
<td>P1</td>
<td>5</td>
<td>6*</td>
<td>.61</td>
<td>Standard</td>
</tr>
<tr>
<td>P20</td>
<td>5</td>
<td>6*</td>
<td>.64</td>
<td>Standard</td>
</tr>
<tr>
<td>P15</td>
<td>5</td>
<td>5</td>
<td>.87</td>
<td>Intralingual</td>
</tr>
<tr>
<td>P17</td>
<td>5</td>
<td>5</td>
<td>.61</td>
<td>Intralingual</td>
</tr>
</tbody>
</table>

*learning effect

Note: 6: correct word, correct spelling. 5: correct word, incorrect spelling.

As can be deduced from table 6 above, both cases in which spelling improvement occurred, participants were exposed to Dutch audio with English subtitles and had spent at least .61 seconds in the AOI of trophy. From table 6 it becomes clear that even though all four participants were exposed to the English subtitles and they spent approximately the same amount of time looking at the AOI of trophy, a learning effect was not visible in each of the cases, but only in the standard subtitling condition.

While trophy is a word that was known by most participants in the pre-test, watery grave show different test scores on both pre- and post-test. Watery grave is pronounced once in each film fragment, and occurs only once for 2.602 (standard, ‘zeemansgraf’ Dutch translation of ‘watery grave’), 2.039 (reversed), or 2.339 (intralingual) seconds in the subtitles. There was only one participant who gave the correct answer for watery grave in both pre- and post-test. Only 12 participants showed a learning effect for this word, as can be seen from table 7 below.

Table 7

Test scores on watery grave and time fixated on AOI. Time is in seconds.

<table>
<thead>
<tr>
<th>Participant</th>
<th>Score pre-test</th>
<th>Score post-test</th>
<th>Time spent in AOI</th>
<th>Condition</th>
</tr>
</thead>
<tbody>
<tr>
<td>P14</td>
<td>0</td>
<td>6*</td>
<td>.47</td>
<td>Reversed</td>
</tr>
<tr>
<td>P5</td>
<td>4</td>
<td>6*</td>
<td>.47</td>
<td>Reversed</td>
</tr>
<tr>
<td>P13</td>
<td>4</td>
<td>6*</td>
<td>1.03</td>
<td>Reversed</td>
</tr>
<tr>
<td>P1</td>
<td>4</td>
<td>4</td>
<td>.57</td>
<td>Reversed</td>
</tr>
<tr>
<td>P7¹</td>
<td>4</td>
<td>4</td>
<td>.35</td>
<td>Reversed</td>
</tr>
<tr>
<td>P20</td>
<td>4</td>
<td>4</td>
<td>.79</td>
<td>Reversed</td>
</tr>
<tr>
<td>P6¹</td>
<td>0</td>
<td>6*</td>
<td>-</td>
<td>Intralingual</td>
</tr>
<tr>
<td>P2¹</td>
<td>4</td>
<td>6*</td>
<td>.38</td>
<td>Intralingual</td>
</tr>
</tbody>
</table>
It shows that a learning effect was visible in 11 out of 20 cases. Only one participant from the control condition showed an improvement. In the reversed condition, participants can only show a learning effect when subtitles are read since that is the only source of target language input. From the data in table 6 it can be concluded that even while having fixated on the AOI, it does not mean that the word is automatically remembered. For example, in the intralingual and standard condition, improvement took place regardless of the fact that they did or did not read the subtitles. In the intralingual condition, learning did not always occur, whereas in the standard condition, aside from the one participant who had already provided the correct answer in the pre-test, learning took place in all four cases.

The third target word that will be looked at more closely, is *inclement*. Whereas the other two words are nouns and are expected to be learned more quickly (d'Ydewalle & van der Poel, 1999, p. 240), this is an adjective. The word is displayed in the subtitles for 3.65 seconds (standard) (‘meedogenloos’ (Dutch translation of ‘inclement’) occurs in the second line of the 2-line subtitle), 1.07 seconds (reversed), or 1.37 seconds (intralingual). Something interesting happened in the data for this target word. As can be seen from table 7 below, there was not a single participant who wrote down the correct target word in the pre-test.
Table 8

*Test scores on inclement and time fixated on AOI. Time is in seconds.*

<table>
<thead>
<tr>
<th>Participant</th>
<th>Score pre-test</th>
<th>Score post-test</th>
<th>Answer post-test</th>
<th>Time spent in AOI</th>
<th>Condition</th>
</tr>
</thead>
<tbody>
<tr>
<td>P14</td>
<td>0</td>
<td>5*</td>
<td>inclement</td>
<td>.66</td>
<td>Reversed</td>
</tr>
<tr>
<td>P1</td>
<td>4</td>
<td>6*</td>
<td></td>
<td>.20</td>
<td>Reversed</td>
</tr>
<tr>
<td>P7¹</td>
<td>4</td>
<td>6*</td>
<td></td>
<td>.69</td>
<td>Reversed</td>
</tr>
<tr>
<td>P5</td>
<td>4</td>
<td>5*</td>
<td>increment</td>
<td>.51</td>
<td>Reversed</td>
</tr>
<tr>
<td>P13</td>
<td>2</td>
<td>4</td>
<td></td>
<td>.19</td>
<td>Reversed</td>
</tr>
<tr>
<td>P20</td>
<td>4</td>
<td>2</td>
<td></td>
<td>.50</td>
<td>Reversed</td>
</tr>
<tr>
<td>P6¹</td>
<td>0</td>
<td>6*</td>
<td></td>
<td>-</td>
<td>Intralingual</td>
</tr>
<tr>
<td>P15</td>
<td>0</td>
<td>5*</td>
<td>inclemation</td>
<td>1.06</td>
<td>Intralingual</td>
</tr>
<tr>
<td>P2¹</td>
<td>4</td>
<td>6*</td>
<td></td>
<td>.09</td>
<td>Intralingual</td>
</tr>
<tr>
<td>P11</td>
<td>4</td>
<td>6*</td>
<td></td>
<td>-</td>
<td>Intralingual</td>
</tr>
<tr>
<td>P17</td>
<td>2</td>
<td>2</td>
<td></td>
<td>.20</td>
<td>Intralingual</td>
</tr>
<tr>
<td>P9</td>
<td>0</td>
<td>6*</td>
<td></td>
<td>-</td>
<td>Control</td>
</tr>
<tr>
<td>P4</td>
<td>4</td>
<td>6*</td>
<td></td>
<td>-</td>
<td>Control</td>
</tr>
<tr>
<td>P16</td>
<td>4</td>
<td>5*</td>
<td>incummulent</td>
<td>-</td>
<td>Control</td>
</tr>
<tr>
<td>P3</td>
<td>0</td>
<td>4</td>
<td></td>
<td>-</td>
<td>Control</td>
</tr>
<tr>
<td>P10</td>
<td>0</td>
<td>5*</td>
<td>impediment</td>
<td>.67</td>
<td>Standard</td>
</tr>
<tr>
<td>P19</td>
<td>2</td>
<td>5*</td>
<td>incremental</td>
<td>.75</td>
<td>Standard</td>
</tr>
<tr>
<td>P18</td>
<td>4</td>
<td>6*</td>
<td></td>
<td>.51</td>
<td>Standard</td>
</tr>
<tr>
<td>P12</td>
<td>4</td>
<td>4</td>
<td></td>
<td>-</td>
<td>Standard</td>
</tr>
<tr>
<td>P8¹</td>
<td>4</td>
<td>2</td>
<td></td>
<td>-</td>
<td>Standard</td>
</tr>
</tbody>
</table>

*Learning effect*

Note: 6 & 5: correct word; 4 & 3: related word; 2 & 1: unrelated word; 0: no response

From table 8 it becomes clear that, even though participants did not know the word in the pre-test, even without being exposed to the correct spelling (such as in the control condition and the standard condition), they provided a correct answer in the post-test. However, there were also six participants who apparently did not know the word beforehand, but did try to use the newly acquired word. In the standard condition, there was only one participant who wrote the word down correctly, whereas the other two participants who showed a learning effect, heard the word, but wrote down a word that only approached the correct spelling. In the reversed and intralingual conditions, there are more participants who used the correct
spelling (two per condition, as opposed to only one in the standard condition), but also two who used an incorrect spelling of the target word. However, it can be deduced that the attempts at the target word are more closely related to the correct spelling in the reversed and intralingual condition (exposure to the correct written form), than are those in the control and standard condition (no exposure to the correct spelling).

5. Discussion
The aim of this study was to find out whether there is a difference in efficacy in the process of FL vocabulary acquisition between three different types of subtitling (i.e. reversed, intralingual, and standard). Two sub-components of the study aimed at finding out whether there is a difference in the number of misspellings of target words between the various subtitling conditions, and whether the time spent reading the subtitles is related to the acquisition of the target vocabulary.

This chapter provides an interpretation of the results displayed in the chapter 4 above. The three research questions will be addressed separately, with the help of results from the current experiment. Results will be explained in the context of previous research. After the research questions have all been addressed, several suggestions are made for further research.

5.1 Learning effect
The analysis of the results should point out whether there was a difference between the effect on FL vocabulary acquisition of the different subtitling conditions (standard, intralingual, reversed, or no subtitles). It was hypothesised that there would be a difference between the groups, and that the participants in the intralingual condition showed the greatest learning effect. This was expected since all participants have a high level of language proficiency (C1-C2 according to the CEFR (Council of Europe, 2001)) and people with this high a level of proficiency do not use their L1 for interpretation of an FL (Weinrich, 1953) The intralingual condition does not make any link to the participants’ L1 and it was therefore expected that test subjects would benefit from this situation.

Results showed that there was a significant learning effect: watching the film fragment had had a positive effect on vocabulary knowledge. However, results do not point toward one condition being significantly more effective than the others. Even though studies have pointed toward the surprising effectiveness of reversed subtitling because, this effect was not found in the current study, nor was the expected positive effect of the intralingual condition on high language proficiency level students visible. In fact, as can be seen from figures 3 and 4 above, the learning effect was of approximately the same size for each of the
conditions. Surprisingly, however, there was also no significant difference between the conditions that did provide subtitles (reversed, intralingual, standard), and the condition that did not (control). What can be established from these results, contrary to what Sweller (1988; 2004) and Zhang (2013) have argued, is that indeed it does not take any extra effort to process subtitles in addition to the soundtrack, as had already been proposed by d’Ydewalle (2002), Kruger et al. (2013), and Navon and Groopher (1979). However, it must be noted that in the present study, only vocabulary was taken into account and no scientific instruments were used to see how active the participants’ were while being exposed to the film fragment (EEG measurements, like Kruger et al. (2013) used), nor was pupil dilation taken into account. Both types of data would have provided insight into the question whether indeed multimodal input takes effort to process (i.e. increase cognitive load). If it were the case that it would take more effort to process bimodally and, as suggested by Sweller (1994) and Zhang (2013), language acquisition would in fact suffer from having to divide one’s attention over different modalities, then the control condition would have put forward significantly higher test scores on the post-tests. Since this was not the case, and all subtitling conditions produced relatively identical results, it can be concluded that it is very well possible to process different modalities at the same time.

However, what cannot be concluded from these data, is that there is one subtitling condition that is more effective in classroom settings than another. There are different factors that play part in the absence of significant results. One evident reason is that the number of participants was very small. In order to find meaningful results, it would have been preferable to have at least twenty participants per condition. In the current study, there were only four to six participants in each condition.

Even though it was expected that the reversed condition showed a great learning effect (as proposed by Danan (1992) and d’Ydewalle (2002)), it did not turn out to be any more effective than the other subtitling conditions. Participants all grew up with Dutch as their native language, and none of them grew up in Belgium or with a southern accent. Participants may therefore have found it disturbing that the reversed condition contained Flemish accents in the soundtrack: it may have been that participants paid more attention to the way the Dutch words were pronounced, than at their English translations in the subtitles.

Another plausible reason that no significant difference were found between the four conditions is that even though participants were asked to leave no blanks in neither pre- nor post-test, not everyone took due account of this instruction. The fact that many students left blank spaces in the pre-test and did less so in the post-test, means that a learning effect that has been administered, is not solely because of the film fragment, but may also have been because of the fact that in the post-test it became more clear to participants that they...
were expected to leave no open spaces. While figure 5 shows that differences between number of blanks in pre- and post-test are no greater than 4, there were participants who showed a difference of up to 10. Some showed a decrease of 5, some of 4, and some showed an increase of 1. This does, however, mean that results are difficult to interpret: did participants leave blanks open because of lack of time during the pre-test? Perhaps the familiarity with the target words during the post-test left participants with more time to think of the correct answers to fill the spaces that they earlier left blank. There was, however, no feedback session after the experiment was completed, so no information is available on why participants left blank spaces.

A final plausible explanation for the lack of significant results, is that practically all participants mentioned that they were expecting the final set of questions to be the same questions as they received in the pre-test. They said they noticed this rather soon after the film fragment had started. They explained that they tried to find the words in the film fragment that they did not know before. This way, the priming on the target words was especially obvious. In an experiment testing incidental vocabulary acquisition, this would have been a complication: it would no longer be incidental because participants had been primed to listen for the words they were expected to know. However, when the question to answer is whether film provides for a good opportunity to have students encounter new words and be aware of the fact that they do so (Rutherford, 1987; Smith, 1991), causing students to notice new words and to stimulate students to remember them, is not such a bad strategy after all. It seems that, when students are triggered to notice new vocabulary, it does not matter which subtitling condition they are exposed to: they remember at least a few items of the FL vocabulary. A problem in how this theory relates to the present study, is that the target vocabulary proved to be not very new to the participants. Participants were of a relatively high language proficiency level and the chance that they had already encountered the vocabulary beforehand is big.

5.2 Misspellings of target words
The question whether more misspellings of target vocabulary occurred in the post-test in the conditions where subtitles were either non-existent, or in the L1, can be answered negatively. It was hypothesised beforehand that more errors would be made in newly acquired vocabulary when participants had not been exposed to the correct spelling in the subtitles. When, on the other hand, participants were exposed to the correct spelling, less mistakes would be made in the spelling of the target vocabulary. On the whole, it was expected that more errors would be made in the post-test than in the pre-test since
participants would probably use familiar words in the pre-test and it is unlikely that familiar words are written incorrectly.

The data showed no significant differences between the four groups. Even though figure 6 on page 33 seems to imply that there is a greater difference between the number of misspellings in the control and standard conditions, it is the number of misspellings in the post-test that count: these are the misspellings that are influenced by the film fragment. It can be seen from figure 6 that the intralingual condition is at approximately the same level with the control and standard conditions as far as misspellings in the post-test are concerned. Furthermore, the differences between the number of misspellings are very small: the increase of misspellings lies between 0.5 and 1.5, meaning that the difference between the conditions is only one misspelled word. When one takes into account that there were twenty words that could have been misspelled, one word accounts for only 5% of the cases.

Even though it is highly likely that completely unknown words are spelled incorrectly after having received input that was solely auditory, it is possible that the target words in this study (table 4 on page 24) are not unheard of by participants of an advanced proficiency level. Furthermore, given the experience that most advanced level participants have in using English, it can be expected that they have the capability to correctly write words that have heard only once. The fact that no significant differences are found here, is therefore not surprising. It would therefore be interesting for further studies to include not only high level proficiency students, but also lower level proficiency students. It may very well be the case that there is a difference in processing and therefore efficacy of subtitles (Weinrich, 1953).

5.3 Eye tracking
The eye tracking component was added to the experiment to see whether there would be any additional information on if participant looked at the subtitles, and if so, whether they looked at it long enough to remember the target words. There have been several studies that have incorporated eye tracking into their experiments (Bisson et al., 2012; d’Ydewalle and de Bruycker, 2007) but the results were inconclusive.

In this study, the most important question that was attempted to be answered, is whether the more time spent in the subtitled area lead to a greater learning effect. However, since there were no significant differences in learning effects between the different subtitling conditions, this question is more or less redundant. But that does not mean that information from the eye tracking data cannot contribute to the question which subtitling condition is the most effective. In fact, combining the eye tracking data to the data from the vocabulary tests, provides valuable insight into this very question. For instance, when looking at the
target word *trophy*, it becomes clear that having TL subtitles is very effective. While *trophy* is a word that is known to almost each participant, because of L1 interference, there were five occurrences in which participants wrote ‘trophee’ (from the Dutch ‘trofee’) instead of ‘trophy’. As can be seen from table 5 above, those participants who incorrectly wrote ‘trophee’ changed this to ‘trophy’ in the post-test, after having been exposed to the correct form in the FL subtitles. This, however, only occurred in the reversed condition, and not in the ENEN condition. An explanation may be that in the reversed condition, the difference between the L1 spelling and the FL spelling became clear, after which they changed their spelling. This difference may not have been as obvious for the intralingual condition, since the L1 is not present there. However, this is just a speculation since no research has been done on this effect yet. It may be interesting for further research to see whether this is an effect that occurs more often, or that it is just a coincidence that this happened.

The second word that was looked at more closely, was *watery grave*. A learning effect was visible in all four conditions, but it was the most obvious in the subtitled ones. In the unsubtitled condition (control), there was only one participant who showed a learning effect, whereas in the other three conditions, there were either three or four participants who used the correct target word in the post-test and did not do so in the pre-test. This shows that even though learning can take place without subtitles, it is more likely to occur when subtitles are present. No relation has been found between the time participants fixated on the target word in the subtitles and whether or not they showed a learning effect. Furthermore, the eye tracker was not able to track eye movements 100% of the time. Part of the reason why 100% was not reached, is because participants blink. At least 80% of the eye movements should have been followed. However, this was often not the case, and in a few cases only 2 to 20 percent of the eye movements was tracked. Therefore the fixation times do not provide very valuable information. What the limited eye tracking data do show, however, is that participants need not look at subtitles for more than one second to be able to see and remember the correct spelling of a target word.

The third and final word that was discussed in greater detail is *inclement*. Inclement is not a well-known word and was therefore not once correctly provided as an answer in the pre-test. However, in the post-test, there were a great number of participants who show to have either heard or read the word in the video. As was presented in the results section, it shows that there are several participants who have used the correct spelling of the word, but that there are also some participants who only came close to the correct spelling. It was concluded that spellings that approximated the correct spelling the most, were found in the FL subtitling condition, whereas the condition without subtitles, or where the subtitles were in the native language, produced spellings that were close to the correct spelling, but that
showed greater deviations than the condition in which FL subtitles were available. Even though the results are not very convincing considering the fact that there is such a small number of participants who show this effect\(^3\), it becomes clear from this small experiment that there are differences in the misspellings. Statistics do not prove this, but when looking at the types of misspellings that were made, it comes to light that participants who have been exposed to the correct spelling, approximate the correct spelling more so than those who have not been exposed to the proper spelling.

The target word *inclement* is remembered by most of the participants, and no differences between subtitling conditions are to be found. The reason why so many participants picked up on this word is probably that the description in the vocabulary test makes it clear that it is a word that has to do with the weather. In the film fragment, there is only one scene in which the weather is discussed and the word is pronounced with emphasis. Even when the word or translation occurs in the subtitles, there are no other words visible at the same time (although this does not hold true for the standard condition). The viewer’s attention is therefore drawn to the word, more so than to *watery grave*, a word that appears in the middle in a sentence and is pronounced without emphasis. Seeing as in some cases learning took place when a word occurred only once, film indeed proves to be a useful medium for implicit language education: new vocabulary is being picked up without learning dictionary definitions (Ellis, 1994; Nation & Waring, 1997; Ponniah, 2011). It may well have been that participants did not pick up on words like inclement, or watery grave, had it not been for the fact that they were primed to listen or look for the target vocabulary in the film fragment. Noticing, as discussed earlier, is one of the main factors in FL vocabulary acquisition and causes participants to more quickly memorise vocabulary (Schmidt, 2010; Smith, 1991).

### 5.4 Suggestions for further research

Due to the fact that no significant results were found with regard to the question whether there is one subtitling condition that is more suitable for language education, this study does not provide a conclusive answer to the question. The number of participants was too small for the study to stand on its own. It is therefore suggested that more research be done with more participants. Furthermore, it is important that a film or film fragment is chosen in which more unknown words occur and that they occur more frequently than once or twice. It is a known fact that the frequency at which someone encounters a word is a crucial factor in the vocabulary acquisition process (Elley, 1989). Not only choosing a film that has many

\(^3\) This presumably is because the word does not occur frequently enough (Elley, 1989).
unknown words in it, but also choosing a film with which it is possible to provide the native language without an accent, is important to reduce distraction and reduce the chance that viewers become annoyed (Bisson et al., 2012). Even though the Flemish accent is basically the same as Dutch, when showing a film in Flemish to Dutch participants, the participants may have to put in an effort to get used to the way the actors speak.

Another idea to overcome language problems and problems in administering learning effects – instead of seeing an increase of knowledge as learning, whereas it actually only means that in the pre-test a word could not be conjured up and that during the post-test this word suddenly was retrieved – it would be a good choice to choose Esperanto as a foreign language. Esperanto is a fantasy language, like Elfish in *The Lord of the Rings* but more easy to acquire, and that nobody really speaks. When a learning effect occurs in a situation where Esperanto is used as the FL, this can only be interpreted as learning, not as previous knowledge coming back to mind since there is no previous knowledge. It would, however, take precious time and a great amount of effort to prepare the materials, since the soundtracks and the subtitles will need to be remade. On top of that, it may prove a challenge to find someone who speaks Esperanto well enough to dub the film fragment.

It may furthermore be the case that beginning learners of a language desire a different type of subtitling than advanced learners of a language (Weinrich, 1953; Lambert and Holobow, 1984). Because of the different ways in which the various groups of learners process a language, advanced learners presumably benefit more from intralingual subtitling, whereas beginner learners access an L2 with the help of their L1, in which case standard or reversed subtitling would do more good. More research is needed to find out whether there is truth in this assumption. When research has pointed out that one condition is better for beginner groups of language learners, and another for a more advanced group, teachers can base their choices in the classroom on these results. This would, in turn, lead to better language education.

On top of the language issues, when eye tracking is incorporated into a study, it should be made sure that the eye tracker is able to record all, or at least 80%, of the eye movements. While the current technology does not allow for a 100% recording yet – it is normal that 20% of the test subjects is not recorded properly – the problem may be solved when improvements to the eye-tracker itself are made. The best that can be done by the researcher himself or herself is to make sure that the eye tracker is properly calibrated and adjusted to the participants’ sitting positions. When less than 50% of the eye movements has been recorded, one had better not use the results from the eye tracker since it may then very well be the case that the results are incomplete, or that the participant in question had
actually looked longer in a particular area, but that the extra time was not recorded. Drawing conclusions from incomplete data is not safe.

6. Conclusion
The present study aimed to contribute to the discussion about subtitling and its usefulness with regard to foreign language vocabulary acquisition. Given the many different views and limited research that has been conducted that have included all types of subtitling available, there is a call for more research. Many studies have already established that subtitling is a medium to enhance understanding, and to transform incomprehensible input into comprehensible input. However, subtitling is more than simply providing a translation of what is being said: there are different types of subtitles, each of which has different implications to the process of FL vocabulary acquisition. Reaching conclusive answers to the question which subtitling condition is the most effective is especially important for language education, since film has gained popularity in this situation and has taken over the role of providing authentic input for language learners.

To try and provide a conclusive answer which subtitling condition best to use in the language classroom, an experiment was set up that examined the effect of the different subtitling conditions on acquisition of target words. This was done with the help of a pre-test, a film fragment, and a post-test. Results from the two vocabulary tests were compared to see whether there was a significant difference. There was: the film fragment caused the desired effect. However, the expected effect of significantly higher test-scores in one of the subtitling conditions, was not found. Instead, each condition had approximately the same effect on the test-scores, even the condition in which no subtitles were provided at all. This finding does away with the idea that providing subtitles in any form has a greater positive effect than providing no subtitles. The proposed explanation for the lack of significant results is that the number of participants in the present study is very small. In order to find any significant results, larger numbers of participants need to be tested. It is suggested that more research is done, since there are many studies that point toward the likelihood that there is indeed a difference in effect on vocabulary acquisition between the different subtitling conditions (e.g. d’Ydewalle & de Bruycker, 2007; Danan, 1992; Lambert & Holobow, 1984). It would therefore be worthwhile to set up a similar experiment and to include a greater number of participants, perhaps even participants of various levels of FL proficiency.

Even though no large effects were found, some small differences were found that show hope that in experiments with more participants, some significant results must be found. One of these is the number of misspellings made when a newly acquired word is used
for the first time. It was found that when participants only received auditory FL input, attempts at using the new word were not as close to the correct spelling of the word as the attempt by participants who had received the written form of the target word in the input. Even though this effect was only found in one of the target words because most of the target words were already known by the participants before the experiment started (or so it seemed), in a similar experiment with more unknown words, this effect may be significant.

Therefore, I strongly suggest that more research be done in this area, so that teachers who want to use film in the language classroom in the most effective way possible, can base their decisions on research pointing towards the different effects of the three subtitling conditions. This way, they can choose the subtitling conditions that best fits the students’ needs.
Bibliography


Appendix A
Questionnaire

Persoonlijke informatie
Naam: .................................................................................................
E-mail adres: ............................................................................................
Geslacht: □ Man □ Vrouw
Geboortedatum: ..........................................................................................
Geboorteplaats: ..........................................................................................
Wat is je hoogst voltooide opleiding?
□ Middelbare school □ MBO
□ HBO □ Universiteit
Huidige studie: ..........................................................................................
Studiejaar: .................................................................................................

Vragen over taalvaardigheid
Heb je ooit in andere landen gewoond?
□ Ja □ Nee
Zo ja, waar en hoe lang?
..........................................................................................................................
..........................................................................................................................
..........................................................................................................................

Welke talen heb je geleerd voordat je naar school ging?
□ Nederlands
□ een andere taal/andere talen namelijk
..........................................................................................................................
..........................................................................................................................

Welke talen heb je op school of op werk geleerd? Hoe vaardig ben je in de verschillende talen? Geef een cijfer voor elke vaardigheid:
1 = erg slecht, 2 = slecht, 3 = gemiddeld, 4 = goed, 5 = heel goed
Taal: .............................................................Lezen:......Schrijven:......Spreken:......Luisteren:......
Taal: .............................................................Lezen:......Schrijven:......Spreken:......Luisteren:......
Taal: .............................................................Lezen:......Schrijven:......Spreken:......Luisteren:......
Taal: .............................................................Lezen:......Schrijven:......Spreken:......Luisteren:......

Welke talen heb je buiten school of werk geleerd? Hoe vaardig ben je in de verschillende talen? Geef een cijfer voor elke vaardigheid:
1 = erg slecht, 2 = slecht, 3 = gemiddeld, 4 = goed, 5 = heel goed
Taal: .............................................................Lezen:......Schrijven:......Spreken:......Luisteren:......
Taal: .............................................................Lezen:......Schrijven:......Spreken:......Luisteren:......
Taal: .............................................................Lezen:......Schrijven:......Spreken:......Luisteren:......
Taal: .............................................................Lezen:......Schrijven:......Spreken:......Luisteren:......

Vragen over taalgebruik
Op een schaal van 1 tot 5 (zie hierboven), hoe goed spreek je Engels? ..... 
Op een schaal van 1 tot 5 (zie hierboven), hoe goed spreek je Nederlands? .....
Hoe vaak spreek je Nederlands?
- Bijna nooit
- Een paar keer per jaar
- Maandelijks
- Wekelijks
- Dagelijks

Hoe vaak spreek je Engels?
- Bijna nooit
- Een paar keer per jaar
- Maandelijks
- Wekelijks
- Dagelijks

Hoeveel procent van je vrienden is Nederlands? ..........%
Hoeveel procent van je vrienden is Engels? ..........%

Welke taal spreek je het makkelijkst?
- Engels
- Nederlands
- Geen verschil

Hoeveel procent van de tijd spreek je Engels met familieleden? ..........%
Hoeveel procent van de tijd spreek je Engels met vrienden? ..........%
Hoeveel procent van de tijd spreek je Engels met collega’s? ..........%
Hoeveel procent van de tijd spreek je Engels als je niet thuis bent? ..........%
Hoeveel procent van de media (film etc.) is in het Engels? ..........%
Hoeveel procent van wat je leest (boeken, kranten) is in het Engels? ..........%
Appendix B

Adjusted subtitles

B.1 Standard

1 00:00:10,106 --> 00:00:13,474 Luister, druiloren, we laten ze iets van piraterij zien.
2 00:00:15,917 --> 00:00:20,578 Er is nog tijd. We vullen ze met glinsterende buit.
3 00:00:20,612 --> 00:00:23,856 Ik kan die trofee al proeven.
4 00:00:23,891 --> 00:00:26,761 Trofee.
5 00:00:30,600 --> 00:00:33,099 Kapitein! Een zeil in zicht.
6 00:00:34,081 --> 00:00:37,021 Er achteraan, Nummer Twee.
7 00:00:37,056 --> 00:00:41,447 Hijs de zeilen.
8 00:01:03,913 --> 00:01:07,055 Sta klaar bij de kanonnen, mannen.
9 00:01:07,325 --> 00:01:10,974 Vuur kanonnen vier en zes af.
10 00:01:17,460 --> 00:01:20,028 Eet kanonskogels, zwabbers.
11 00:01:23,913 --> 00:01:26,988 Observeer en leer.
12 00:01:33,000 --> 00:01:35,061 Ik ben de Piraten Kapitein.
13 00:01:35,900 --> 00:01:40,799 En ik ben hier voor jullie goud.
14 00:01:41,009 --> 00:01:43,440 's Werelds beste kapitein.
15 00:01:43,475 --> 00:01:45,196 's Werelds beste kapitein.
16 00:01:46,751 --> 00:01:51,785 's Werelds beste kapitein.
17 00:01:53,744 --> 00:01:57,595 's Werelds beste kapitein.
18 00:01:58,206 --> 00:01:59,742 's Werelds beste kapitein.
19 00:02:02,798 --> 00:02:07,292 's Werelds beste kapitein.
20 00:02:08,305 --> 00:02:10,940 's Werelds beste kapitein.
21 00:02:11,379 --> 00:02:14,319 Spijkerbroek.
22 00:02:19,420 --> 00:02:20,738 Sorry.
23 00:02:51,549 --> 00:02:55,636 Sorry.
24 00:03:00,780 --> 00:03:03,279 's Werelds beste kapitein.
25 00:03:03,314 --> 00:03:05,509 's Werelds beste kapitein.
26 00:03:06,962 --> 00:03:08,516 's Werelds beste kapitein.
27 00:03:13,826 --> 00:03:17,811 's Werelds beste kapitein.
28 00:03:18,246 --> 00:03:22,908 's Werelds beste kapitein.
29 00:03:22,909 --> 00:03:24,817 's Werelds beste kapitein.
30 00:03:24,843 --> 00:03:28,545 's Werelds beste kapitein.
31 00:03:28,787 --> 00:03:33,753 's Werelds beste kapitein.
Maar dat kan niet. Ik bedoel...
De bemanning kunnen niet zonder u.
Wat dacht ik ook, Nummer 2?
Piraat van het Jaar?
Ik?
Kijk nu naar de prijzenkast.
De beste anekdote over een pijnlinkvis.
Dat is al het bewijs van mijn hele carrière.
Maar, maar, Piraat van het Jaar is maar commerciële onzin.
Echte piraterij draait niet om trofeeën.
Maar om achteruit op trappen te vechten.
Glanzende baarden.
En denk aan al de avonturen die we hebben meegemaakt.
Ons avontuur met Aztek.
Ons avontuur met dat varken.
Ons avontuur met...
Sst!
Oké, misschien niet met de bibliothecaressen, maar wat 'n plezier.
En wat dacht je van de zeemansliedjes?
Extra griezelig.
Dat ze het maar in hun broek doen.
66
00:05:38,425 --> 00:05:41,071
Dagboek van Charles Darwin.
67
00:05:41,206 --> 00:05:44,582
Dag 93 aan boord van de Beagle.
68
00:05:45,417 --> 00:05:50,109
Ik heb vandaag een nieuw soort mossel ontdekt...
69
00:05:50,744 --> 00:05:56,402
die ik in de orde van Pygophora gecategoriseerd heb.
70
00:05:59,677 --> 00:06:03,498
Ik krijg nooit een vriendin.
Ik ben zó ongelukkig.
71
00:06:17,947 --> 00:06:22,609
En nu word ik aangevallen door piraten.
72
00:06:36,021 --> 00:06:40,288
Halt! Ik ben de Piraten Kapitein en ik ben hier voor j e goud.
73
00:06:40,756 --> 00:06:44,239
Ik heb geen goud.
- Geen goud, hè?
74
00:06:44,573 --> 00:06:50,046
Wat is dit dan, als ik mag vragen?
- Dat is een bavianen nier.
75
00:06:51,333 --> 00:06:56,907
En dit dan?
- Dat is nog een bavianen nier.
76
00:06:58,663 --> 00:07:01,419
We hebben het ruim doorzocht, kapitein.
- En?
77
00:07:01,439 --> 00:07:04,176
Alleen maar wezens, stukjes van...
78
00:07:04,244 --> 00:07:08,534
wezens in potten.
- En een trieste baviaan.
79
00:07:11,473 --> 00:07:14,582
Wat voor soort schip is dit?
80
00:07:14,683 --> 00:07:19,818
We zijn, nou ja, we waren, een wetenschappelijke expeditie.
Het weer van vandaag:
meedogenloos.

00:08:35,325 --> 00:08:38,906
Stop.
- Wat nu weer? Een laatste verzoek?

00:08:38,974 --> 00:08:43,230
Moeten we die inwilligen? Staan we bij een mensenrechtenconventie ingeschreven?

00:08:43,264 --> 00:08:46,440
Die vogel.
- Mijn papegaai? Ze heeft gewoon zware botten.

00:08:46,542 --> 00:08:49,650
Ik bedoel, ze is geen papegaai.

00:08:49,684 --> 00:08:52,454
Geen papegaai?
Waar heeft hij het nou weer over, meisje?

00:08:52,521 --> 00:08:57,927
Ze is de wetenschappelijke ontdekking van deze tijd. Ze is een...

00:09:00,629 --> 00:09:03,096
Sorry.
Het is gewoon...

00:09:03,197 --> 00:09:05,900
Dat is mijn favoriete gedeelte.

00:09:10,900 --> 00:09:15,528
Een dodo, zeg je?
- Ze zijn al 150 jaar uitgestorven.

00:09:15,596 --> 00:09:19,312
Om er tegenwoordig één te vinden die nog leeft, is behoorlijk onvoorstelbaar.

00:09:19,379 --> 00:09:24,002
Slimme meid. Je zal niet uitsterven.
- Als ik zo vrijpostig mag zijn, ben ik bereid...

00:09:25,093 --> 00:09:28,708
om je 10 pond te geven voor jouw Polly.

00:09:28,809 --> 00:09:32,086
Polly is niet te koop.
- 10 pond.

00:09:32,187 --> 00:09:35,972
Ze hoort bij de familie.
- Ze is als een tante, met een snavel.
- Ik vind gevaar helemaal niet leuk. 127
00:10:43,540 --> 00:10:46,923
Kapitein, kunt u zich ons gesprekje nog herinneren?
128
00:10:47,058 --> 00:10:50,403
Het gesprek waarin in ging over of varkens eigenlijk een soort fruit zijn?
129
00:10:51,112 --> 00:10:52,836
Nee.
130
00:10:56,349 --> 00:11:01,585
Nee, het gesprek over het uit de weg gaan van wilde plannen die eindigen in de dood.
131
00:11:01,653 --> 00:11:06,586
Je hebt de man gehoord, ongekende rijkdom. Dit is het plan:
132
00:11:06,653 --> 00:11:10,268
We gaan naar Londen, Polly wint dat wetenschapsshow dingetje...
133
00:11:10,336 --> 00:11:14,862
we pakken het prijzengeld, ik doe mee aan Piraat van het Jaar, en ik win. Bingo.
134
00:11:14,896 --> 00:11:17,667
Maar kapitein, ik...
- Dit zal niet mislukken.
135
00:11:17,735 --> 00:11:21,316
En dan nog, kijk eens naar haar kleine gezichtje...
136
00:11:21,383 --> 00:11:27,228
Ik wil die wetenschapsprijs zeker wel winnen.
- Bravo, kapitein.
137
00:11:27,329 --> 00:11:31,079
Ik zie dat u een man met visie bent.
138
00:11:32,194 --> 00:11:36,620
En Charles, wanneer is die wetenschap show?
- Morgen over een week.
139
00:11:36,754 --> 00:11:38,545
Verdorie!
140
00:11:38,546 --> 00:11:40,483
We hadden het kunnen redden met een goede wind achter ons.
141
00:11:40,484 --> 00:11:43,469
maar jammer genoeg zit er een vies, groot zeemonster in de weg.
142
00:11:44,795 --> 00:11:50,707
Volgens mij zetten ze die gewoon op de map voor de versiering, kapitein.
143
00:11:52,836 --> 00:11:54,694
Is dat een feit?
144
00:11:54,796 --> 00:11:56,130
Nou, val nu omver.
145
00:11:57,332 --> 00:12:00,575
Je kan nog iets leren van deze toffe kerel, Nummer Twee.
Reversed

1 00:00:08,744 --> 00:00:10,553 All right, now listen up, you coves.
2 00:00:10,612 --> 00:00:14,122 We'll show those swabs a thing or two about pirating.
3 00:00:14,516 --> 00:00:19,261 There's still time to fill these chests with sparkling booty.
4 00:00:19,321 --> 00:00:21,961 I can practically taste that trophy!
5 00:00:22,057 --> 00:00:24,628 Trophy! Trophy! Trophy!
6 00:00:28,798 --> 00:00:31,109 Captain! Sail off the port bow!
7 00:00:32,768 --> 00:00:34,208 Let's get after her, Number Two.
8 00:00:34,269 --> 00:00:35,314 Aye, aye, sir.
9 00:00:35,337 --> 00:00:36,680 Clap on all sail!
10 00:00:37,239 --> 00:00:39,549 Royals and topgallants!
11 00:00:39,608 --> 00:00:40,609 Release those gallants!
12 00:01:00,865 --> 00:01:03,539 Fire those long things that go bang.
13 00:01:04,501 --> 00:01:07,607 Fire cannons four and six!
14 00:01:14,211 --> 00:01:16,018 Eat cannonball, lubbers.
15 00:01:19,984 --> 00:01:22,860 Here we go, lads! Look and learn.
16 00:01:25,289 --> 00:01:27,330 Go get 'em, Pirate Captain!
17 00:01:27,925 --> 00:01:30,906 Avast! I'm the Pirate Captain.
18 00:01:31,296 --> 00:01:33,239 And I'm here for your gold.
19 00:01:33,331 --> 00:01:37,973

Gold? This is a plague boat. I'd give my right arm for some gold.
20 00:01:39,004 --> 00:01:40,506 Or my left.
21 00:01:41,172 --> 00:01:43,678 Right, lads. Minor hiccup.
22 00:01:43,742 --> 00:01:46,387 This time, it's payday!
23 00:01:48,147 --> 00:01:51,920 I'm the Pirate Captain, and I'm here for your gold.
24 00:01:52,351 --> 00:01:54,456 We're on a field trip.
25 00:01:56,990 --> 00:01:59,067 I'm the Pirate Captain, and...
26 00:01:59,825 --> 00:02:01,168 Naturists.
27 00:02:02,194 --> 00:02:05,698 Pirate Captain. Blah, blah, blah, blah, gold.
28 00:02:05,764 --> 00:02:07,141 Ghost ship.
29 00:02:12,372 --> 00:02:13,817 Sorry.
30 00:02:44,422 --> 00:02:46,422 World's best Pirate...
31 00:02:53,248 --> 00:02:57,321 Captain, we've sighted another ship, sir.
32 00:02:58,354 --> 00:02:59,992 That's nice.
33 00:03:05,100 --> 00:03:06,835 Do you want to give the order to attack?
34 00:03:06,900 --> 00:03:11,786 No, not really, Number Two.
35 00:03:12,080 --> 00:03:13,823 No, I think... I think I'll quit piracy.
36 00:03:14,260 --> 00:03:16,035 I'm hanging up my cutlass.
37
Yeah, I was thinking
I might go into baby clothes.

What?

It seems there's a fortune to be made
in baby clothes because
babies grow so fast.

No, but you can't. I mean, the crew...

They'd... They'd be lost without you.

What was I thinking, Number Two?

Pirate of the Year?

Look at my trophy cabinet.

Second Best Anecdote About a Squid.

I mean, that is all I've got to show
for my entire career.

But... Pirate of the Year?

It's just commercialized nonsense.

Real piracy isn't about trophies.

It's about
fighting up staircases backwards.

It's about sliding down sails
with a knife in your teeth.

Beard glossiness!

And think about
all the adventures we've had.

Our adventure with Aztecs.

Our adventure with that pig.

Our adventure in...

Maybe not the library.

But the fun we've had.

And what about the shanties?

People would miss my shanties.

And the running people through.

Enjoyable? You're brilliant at it!

You're a brine-soaked terror
of the high seas!

Do you think so, Number Two?

Everyone does!

You're more of a pirate
than Black Bellamy or any of them.

You're a real pirate.

That's why the crew,

they think you're the best thing
since boil-in-the-bag ham.
Why do you think they all got that tattoo? Come on.
00:04:49,420 --> 00:04:51,324
What do you say?
00:04:51,710 --> 00:04:53,185
Arrr!
00:04:53,200 --> 00:04:55,703
By Neptune, where's that ship?
00:04:57,400 --> 00:04:59,038
This is the one, lads!
00:04:59,300 --> 00:05:02,579
Ninth time lucky!
I can feel it in my beard!
00:05:05,259 --> 00:05:04,699
Well, don't just sit there like a bag of lemons!
00:05:04,780 --> 00:05:06,088
Hoist the flag!
00:05:06,460 --> 00:05:07,768
Standard flag?
00:05:07,840 --> 00:05:09,747
Or extra gruesome?
00:05:09,830 --> 00:05:11,173
Extra gruesome!
00:05:11,180 --> 00:05:13,285
Let's make their gizzards shake.
00:05:24,000 --> 00:05:26,503
Journal of Charles Darwin.
00:05:27,070 --> 00:05:30,482
Day 93 aboard the Beagle.
00:05:30,680 --> 00:05:35,459
I have, today, discovered a new kind of barnacle.
00:05:35,470 --> 00:05:41,352
which I have categorized in the order of Pygophora.
00:05:44,600 --> 00:05:48,104
I'll never get a girlfriend.
00:05:53,200 --> 00:05:55,703
By Neptune, where's that ship?
This is the one, lads!
00:04:57,400 --> 00:04:59,038
Ninth time lucky!
I can feel it in my beard!
00:05:05,259 --> 00:05:04,699
Well, don't just sit there like a bag of lemons!
00:05:04,780 --> 00:05:06,088
Hoist the flag!
00:05:06,460 --> 00:05:07,768
Standard flag?
00:05:07,840 --> 00:05:09,747
Or extra gruesome?
00:05:09,830 --> 00:05:11,173
Extra gruesome!
00:05:11,180 --> 00:05:13,285
Let's make their gizzards shake.
00:05:24,000 --> 00:05:26,503
Journal of Charles Darwin.
00:05:27,070 --> 00:05:30,482
Day 93 aboard the Beagle.
00:05:30,680 --> 00:05:35,459
I have, today, discovered a new kind of barnacle.
00:05:35,470 --> 00:05:41,352
which I have categorized in the order of Pygophora.
00:05:44,600 --> 00:05:48,104
I'll never get a girlfriend.
00:06:02,040 --> 00:06:06,182
And now I'm being attacked by pirates.
I just want one tiny bit of success!

One teensy weensy bit of respect from my peers just once in my life!

Is that such a crime?

But you try telling that to the universe.

No. Just once, just once for things to go right.

Not a parrot?

So, are we done here?

Dear Diary, about to meet a watery grave.

It's nothing personal, you understand.

It's just... It's been a tough week, and a good plank walk usually cheers him up.

Will die without reaching second base with a lady.

Get on with it!

Today's weather:

No. Is there some sort of human rights convention we're signed up to?

She's just big-boned!

No, she's not. No, she's not. Not a parrot?

What's he on about?

She's... She's... She's the scientific discovery of our age!

She's a...

Sorry. It's just, that's my favorite bit, actually.

A dodo, you say?

They've been extinct for 150 years.
To find one alive today, it's quite incredible.

Well, clever old girl. Not going extinct.

If I might be so bold, I would be prepared to pay you 10 pounds for your Polly.

Polly's not for sale.

Ten pound.

She's one of the family.

She's like an auntie.

With a beak.

Yeah, I'm afraid they're right, Charles.

Polly here is the feathery heart and soul of the boat.

Come on, little dodo.

It's such a missed opportunity.

If I could've presented her in London to the Royal Society at their big science show...

Very nice, but no question of that.

- She'd have been an absolute sensation.

And of course a shoo-in for the top prize.
The one about us trying to avoid harebrained schemes that end in us facing certain death. You heard the man, Number Two. Untold riches. We go to London, Polly wins this science show thingy, we take the prize money, I enter Pirate of the Year, I win! Bingo! - But, Captain, I'm not sure... - But I am. And besides, look at her little face. I sure want to win that science prize. Unplanned riches. Here's the plan: We go to London, Polly wins this science show thingy, we take the prize money, I enter Pirate of the Year, I win! Bingo! - But, Captain, I'm not sure... - But I am. And besides, look at her little face. I sure want to win that science prize. Squawk! Bravo, Captain! I see you're a man of vision! Right, then, Charles. When is this science show of yours? A week tomorrow. Confound it! You see, we could've made that with a good wind behind us, but unfortunately there's this dirty great sea monster in the way. I think they just add those onto maps for decoration, Captain. Is that a fact? Well, blow me down. Number Two, you could learn something from this fine fellow.
All right, now listen up, you coves.

We'll show those swabs a thing or two about piracy.

There's still time to fill these chests with sparkling booty.

I can practically taste that trophy!

Trophy! Trophy! Trophy!

Captain! Sail off the port bow!

Let's get after her, Number Two.

Aye, aye, sir.

Clap on all sail!

Royals and topgallants!

Release those gallants!

Fire those long things that go bang.

Fire cannons four and six!

Eat cannonball, lubbers.

Here we go, lads! Look and learn.

Go get 'em, Pirate Captain!

Avast! I'm the Pirate Captain.

And I'm here for your gold.

Gold? This is a plague boat, old man.

I'd give my right arm for some gold.

Or my left.

Right, lads.

Minor hiccup, that last one.

This time,

it's payday!

I'm the Pirate Captain, and I'm here for your gold.

Geography field trip.

Pirate Captain.

Blah, blah, blah, blah, gold.

Ghost ship.

Sorry.

Captain, we've sighted another ship, sir.

That's nice.

Do you want to give the order to attack?

No, not really, Number Two.

No, I think I've...

Think I had enough of piracy.

I'm hanging up my cutlass.
Captain, no!

Yeah, I was thinking I might go into baby clothes, actually.

What?

I hear there's a fortune to be made in baby clothes because babies grow so fast.

No, but you can't. I mean, the crew... They'd... They'd be lost without you.

What was I thinking, Number Two?

Pirate of the Year?

Look at the trophy cabinet.

I mean, that is all I've got to show for my entire career.

Pirate of the Year?

Me?

Me.

Me.

Pirate of the Year?

Me.

Me.

Me.

They'd... They'd be lost without you.

What was I thinking, Number Two?

Pirate of the Year?

Me?

Me.

Me.

Pirate of the Year?

Me.

Me.

Pirate of the Year?

Me.

Me.

Me.

Pirate of the Year?

Me.

Me.

Pirate of the Year?

Me.

Me.

Pirate of the Year?

Me.

Me.

Pirate of the Year?

Me.

Me.

Pirate of the Year?

Me.

Me.

Pirate of the Year?

Me.

Me.

Pirate of the Year?

Me.

Me.

Pirate of the Year?

Me.

Me.

Pirate of the Year?

Me.

Me.

Pirate of the Year?

Me.

Me.

Pirate of the Year?

Me.

Me.

Pirate of the Year?

Me.

Me.

Pirate of the Year?

Me.

Me.

Pirate of the Year?

Me.

Me.

Pirate of the Year?
since boil-in-the-bag ham.
73
00:04:57,731 --> 00:04:59,239
Why do you think they all
74
00:04:59,366 --> 00:05:01,812
got that tattoo? Come on.
75
00:05:02,102 --> 00:05:03,206
What do you say?
76
00:05:03,971 --> 00:05:05,746
Arree!
77
00:05:05,839 --> 00:05:08,542
By Neptune, where's that ship?
78
00:05:10,044 --> 00:05:11,882
This is the one, lads!
79
00:05:12,246 --> 00:05:15,325
Ninth time lucky!
80
00:05:15,449 --> 00:05:17,658
Well, don't just
81
00:05:18,118 --> 00:05:19,126
sit there like lemons!
82
00:05:19,286 --> 00:05:20,694
Hoist the flag!
83
00:05:20,954 --> 00:05:23,061
Or extra gruesome?
84
00:05:23,123 --> 00:05:24,266
Extra gruesome!
85
00:05:24,358 --> 00:05:26,663
Let's make their gizzards shake.
86
00:05:38,038 --> 00:05:40,441
Journal of Charles Darwin.
87
00:05:40,507 --> 00:05:44,819
Day 93 aboard the Beagle.
88
00:05:44,878 --> 00:05:49,657
I have, today,
discovered a new kind of barnacle.
89
00:05:49,717 --> 00:05:55,699
which I have categorized
in the order of Pygophora.
90
00:05:59,293 --> 00:06:03,397
I'll never get a girlfriend.
91
00:06:17,044 --> 00:06:21,686
I am so unhappy.
92
00:06:35,229 --> 00:06:39,772
And now I'm being attacked by pirates.
93
00:06:40,234 --> 00:06:42,472
Avast! I'm the Pirate Captain,
94
00:06:42,536 --> 00:06:43,972
and I'm here for your gold.
95
00:06:44,037 --> 00:06:46,483
No gold, eh?
96
00:06:46,907 --> 00:06:48,643
Then what, might I ask, is this?
97
00:06:49,009 --> 00:06:50,981
It's a baboon's kidney.
98
00:06:51,111 --> 00:06:53,355
Is it?
99
00:06:53,614 --> 00:06:55,685
No gold, eh?
100
00:06:58,218 --> 00:06:59,456
We searched the hold, Captain.
101
00:06:59,720 --> 00:07:00,960
And?
102
00:07:01,155 --> 00:07:02,834
Just creatures.
103
00:07:02,890 --> 00:07:05,937
Bits of creatures in jars.
104
00:07:05,993 --> 00:07:08,530
And an unhappy-looking baboon.
105
00:07:11,432 --> 00:07:14,034
What kind of ship is this?
106
00:07:14,168 --> 00:07:15,579
We're a...
107
00:07:15,636 --> 00:07:19,513
Well, we were a scientific expedition.
108
00:07:19,573 --> 00:07:21,712
Charles Darwin, at your service, sir.
109
00:07:21,909 --> 00:07:24,747
For pity's sake!

00:07:24,812 --> 00:07:28,054
Is it so much to ask? Is it?

00:07:28,549 --> 00:07:31,423
I just want one tiny bit of success!

00:07:31,485 --> 00:07:35,328
One teensy weensy bit of respect from my peers
just once in my life!

00:07:37,825 --> 00:07:39,436
Is that such a crime?

00:07:39,493 --> 00:07:40,938
No?

00:07:45,466 --> 00:07:46,501
But you try telling that to the universe.

00:07:48,669 --> 00:07:51,973
Just once,
just once for things to go right.

00:07:52,139 --> 00:07:53,379
Just... Just once.

00:07:56,910 --> 00:08:00,319
So, are we done here?

00:08:04,718 --> 00:08:06,629
Dear Diary,

00:08:06,687 --> 00:08:09,026
about to meet a watery grave.

00:08:09,289 --> 00:08:11,196
It's nothing personal, you understand.

00:08:11,258 --> 00:08:13,131
It's just... It's been a tough week,

00:08:13,193 --> 00:08:16,469
and a good plank walk
usually cheers him up.

00:08:16,530 --> 00:08:20,373
Will die without reaching second base

with a lady.

00:08:20,467 --> 00:08:22,745
Get on with it!

00:08:24,972 --> 00:08:26,076
Today's weather:

00:08:26,774 --> 00:08:28,144
inclement.

00:08:35,149 --> 00:08:36,992
- Stop!
- Now what?

00:08:37,050 --> 00:08:39,694
Last request, is it?
Do we have to grant those?

00:08:39,753 --> 00:08:42,700
Is there some sort of human rights convention we're signed up to?

00:08:42,523 --> 00:08:44,332
- But that bird!
- My parrot?

00:08:44,491 --> 00:08:45,868
She's just big-boned!

00:08:45,926 --> 00:08:48,873
No, she's not.
I mean... I mean, she's not a parrot.

00:08:48,929 --> 00:08:50,306
Not a parrot?

00:08:50,364 --> 00:08:51,975
What's he on about, old girl?

00:08:52,032 --> 00:08:52,970
She's... She's...

00:08:53,066 --> 00:08:56,042
She's the scientific discovery of our age!

00:08:56,069 --> 00:08:57,780
She's a...

00:08:59,907 --> 00:09:02,547
Sorry. It's just,

00:09:02,609 --> 00:09:04,646
that's my favorite bit.

00:09:10,184 --> 00:09:11,822
A dodo, you say?

00:09:11,885 --> 00:09:14,661
They've been extinct for 150 years.
To find one alive today, it's quite incredible.
Well, clever old girl. Not going extinct.
If I might be so bold, I would be prepared to pay you 10 pounds for your Polly.
Polly's not for sale.
Ten pound.
She's like an auntie.
With a beak.
Yeah, I'm afraid they're right, Charles.
Polly here is the feathery heart and soul of the boat.
Come on, little dodo.
It's such a missed opportunity.
If I could've presented her
in London to the Royal Society at their annual science show...
No question of that, I'm afraid.
- She'd have been an absolute sensation.
- Yeah, well, can't be helped.
And of course a shoo-in for the top prize.
The one about us trying
to avoid harebrained schemes
that end in us facing certain death.
You heard the man, Number Two.

Untold riches.
Here's the plan.
We go to London,
Polly wins this science show thingy,
we take the prize money,
I enter Pirate of the Year, I win!
- We could've made that with a good wind behind us,
- unfortunately there's this dirty great sea monster in the way.
I think they just add those onto maps for decoration, Captain.

- But, Captain, I'm not sure...
- It can't fail.

And besides, look at her little face.
I sure want to win that science prize.
Appendix C
Pre- and post-test

**Kennis van Engelse woorden**

**Omschrijving**


1. The final resting place of a sailor who died at sea.
2. Treasures that have a value that is beyond description.
3. An object that people use to stay dry when rain is falling down.
4. The prize, such as a cup or plaque, received as a symbol of victory.
5. The man or woman in front of the classroom from whom you learn things.
6. An animal that lives in the sea, moves very fast, has eight arms, produces ink, and is torpedo-shaped.
7. Piece of women’s clothing that hangs freely from the waist down.
8. The songs that sailors and pirates sing to accompany their physical labour.
9. The water that falls down from the sky which is the reason why people put on waterproof coats.
10. A person that enforces the law and gives you fines if you park your car in a spot where it is not allowed.
11. An object that lies on your bed and on which you put your head when you go to sleep.
12. An object that is used to write words down on paper.
13. A publication that is issued daily which contains current affairs, editorials and advertising.
15. A small object that is used to lock or unlock a door or a box.
16. Description of a weather type that is not mild, but rather severe.
17. A building that people live in.
18. People use this animal for transportation. They can ride it, or they can put it in front of a carriage.

19. The space in a ship where boxes and other goods are stored.

20. Another word for to raise or to lift up.

21. The open ocean waters outside the territory of a single country.

22. A word to describe something that is rash, foolish, reckless, or hardly thought through.

23. This grows on your head and regularly needs to be cut in order for it not to grow too long.

24. Agreeing to the fulfilment of a request.

25. The exclamation that is used when a point has been scored in football (soccer).

26. A spotted animal that has an extremely long neck and lives in Africa.

27. During spring and summer these natural objects are seen everywhere. Often produced to sell in bouquets.

28. Four-legged animal that barks and is considered to be a very loyal pet.

29. A slightly curved sword, used by men at sea.

30. An exclamation that is used to damn something when things do not go as planned.

31. A modern, electronic device that can be used for instance for studying, work, and gaming. It consists of at least a screen and a keyboard.

32. A vehicle with four wheels that is used for modern transportation.

33. A word that is used to describe something that has been dipped in salty water.

34. Valuable objects such as jewellery and gold that is plundered by pirates.

35. Something that is used for studying or pleasure. It has a cover and pages, and contains words.

36. A sea animal that lives in a shell and is usually attached to ship bottoms, rocks, or the skin of whales.

37. Monkeys with long noses, short tales and big, red behinds.

38. A term that is used as a command to have someone stop what they are doing.

39. Type of fruit that is fresh, round and coloured in one of the following colours: red, yellow, or green.

40. A short but serious account that tells the story of an event and is told to amuse people.
Appendix D

Answering sheet

Naam:__________________________________

**Antwoordblad.** Vul hier je antwoorden in.

<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>W</td>
<td>21</td>
</tr>
<tr>
<td>2</td>
<td>U</td>
<td>22</td>
</tr>
<tr>
<td>3</td>
<td>U</td>
<td>23</td>
</tr>
<tr>
<td>4</td>
<td>T</td>
<td>24</td>
</tr>
<tr>
<td>5</td>
<td>T</td>
<td>25</td>
</tr>
<tr>
<td>6</td>
<td>S</td>
<td>26</td>
</tr>
<tr>
<td>7</td>
<td>S</td>
<td>27</td>
</tr>
<tr>
<td>8</td>
<td>S</td>
<td>28</td>
</tr>
<tr>
<td>9</td>
<td>R</td>
<td>29</td>
</tr>
<tr>
<td>10</td>
<td>P</td>
<td>30</td>
</tr>
<tr>
<td>11</td>
<td>P</td>
<td>31</td>
</tr>
<tr>
<td>12</td>
<td>P</td>
<td>32</td>
</tr>
<tr>
<td>13</td>
<td>N</td>
<td>33</td>
</tr>
<tr>
<td>14</td>
<td>L</td>
<td>34</td>
</tr>
<tr>
<td>15</td>
<td>K</td>
<td>35</td>
</tr>
<tr>
<td>16</td>
<td>I</td>
<td>36</td>
</tr>
<tr>
<td>17</td>
<td>H</td>
<td>37</td>
</tr>
<tr>
<td>18</td>
<td>H</td>
<td>38</td>
</tr>
<tr>
<td>19</td>
<td>H</td>
<td>39</td>
</tr>
<tr>
<td>20</td>
<td>H</td>
<td>40</td>
</tr>
</tbody>
</table>

Dialang Luisteren:______ Dialang Lezen: ______