

Ingrid Elisabeth Nufsfjord Kvitnes

Subtitles in the Second Language Classroom

An experimental study with Norwegian learners of English

Master's thesis in English Language

Trondheim, spring 2013

Abstract

This study has explored the use of subtitles in second language acquisition for Norwegian learners of English. We used participants from the VG1 level in upper secondary school; the experiment had two experimental groups and a control group who watched the same video clip with or without subtitles. We found that subtitles aided comprehension of plot in the initial round of testing in addition to the participants' vocabulary size, and this was established through the use of a simple comprehension questionnaire. We also tested for long term effects of learning by using a word definition task and a lexical decision task and found that the subtitles were not predictors of performance. The participants' proficiency level was more important as a predictor, through vocabulary size and input interaction. An experiment with back translation for the Norwegian subtitles showed no signs of priming.

Preface

As a teacher, language learning is of great importance to me. To try to understand how learning works is a challenge, but research in the classroom can help expand our knowledge on how students learn language. I teach English to Norwegian students, and I like to show video clips in my classroom as a way to vary my methods and to show my students a visual representation of a topic. Through this study I have read and learnt a lot about the use of subtitles in language learning and it has made me more aware of how I use this in my teaching.

I am very grateful to have been working with this study, with the aid of the Department of Modern Foreign Languages. I would like to thank my supervisors, Mila Vulchanova and Juhani Järvikivi, for valuable help in this process. I would also like to thank my fellow master student, Lisa Maria Aurstad, who has been a great companion in developing, planning and conducting this experiment. Thanks to Sander Kuitert who assisted us when we needed an extra person to carry out the procedure and to the department who made this possible through funding. I must also thank Kristoffer Berntsen for proofreading this thesis. Finally, Håkon Horsgård Watn, thank you for your feedback on this thesis and your continued support through this year.

Ingrid Elisabeth Nufsfjord Kvitnes

Trondheim, May 2013

Contents

1.0 Introduction	7
2. 0 Theory	9
2.1 Second language acquisition	9
2.2 The role of input in SLA	13
2.3 Subtitling of audio visual material and SLA	15
2.4 Visual material: cartoons/animations	20
2.5 Priming	21
3.0 Method	23
3.1 Research questions and the quantitative method.....	23
3.2 Participants.....	24
3.3 Materials.....	26
3.4 Procedure.....	30
3.5 Validity.....	31
4.0 Results	33
4.1 Background	33
4.2 Comprehension questionnaire	34
4.3 Word definition task.....	37
4.4 Lexical decision task	39
5.0 Discussion	41
5.1 Comprehension questionnaire	41
5.2 Word definition task.....	43
5.3 Lexical decision task	46
5.4 General discussion.....	47
6.0 Conclusion.....	51
Bibliography.....	53
Appendices	59

Appendix A – Parent’s consent form	59
Appendix B – Background questionnaire	61
Appendix C – Comprehension questionnaire.....	67
Appendix D – Word definition task	72
Appendix E – Lexical decision task.....	79

1.0 Introduction

The goal of this project was to get a better knowledge of the influence of subtitles on second language acquisition for Norwegian learners of English. As a teacher, every aspect of language that can contribute to the learning process is of interest to me. To revolutionise teaching cannot be a realistic goal for every teacher, but perhaps some steps can be taken to get more out of techniques one already uses. Video clips are often showed in the classroom to give more information to the viewers on a topic, for example to give students a glimpse of another culture. The use of a video clip does not have to be viewed as limited; it can also be used to teach language if the teacher consciously chooses to. As Vanderplank (1990) writes ‘no teacher, no classroom can provide the amount, the quality, the variety of language in interesting, meaningful and informative and often amusing contexts that television can’ (p. 221). What I then wonder is: can subtitles influence the pupils’ learning process? To explore the conditions of this issue I have done research within three groups at upper secondary level, VG1, and shown them the same film clip with different subtitles or without. They were tested on different areas to see if one version would be preferred due to learning facilitation, or if other factors were significant. It is important to note here that the participants for this study were not selected as individuals and subsequently put into three groups. The groups each represent one class of students from the same school and study level, therefore variety between groups and different outliers were not controllable factors that could have been predicted. Initially, before the testing started, the groups were considered to be equal and thus suitable for this experiment.

This study was created after reading the article “Foreign Subtitles Help but Native-Language Subtitles Harm Foreign Speech Perception” by Mitterer and McQueen (2009). We were curious to see if we would find similar or different effects with Norwegian, as Dutch was used as the first language in their article. I stress here that this study is not a replication of Mitterer and McQueen, but merely inspired by it in its initial creation process. The relation between the studies will be commented on below.

Research question: Do subtitles function as predictors for language learning when using audio visual material? Hypothesis 1: Native-Language subtitles will give the best results for the participants, as their age and proficiency level suggest that foreign subtitles will help less in this situation. Hypothesis 2: The students who saw the clip with Norwegian subtitles will

show signs of priming in the lexical decision task, after back translating words from the Norwegian subtitles.

2.0 Theory

2.1 Second language acquisition

Saville-Troike (2006) defines second language acquisition (from here SLA) as “both the study of individuals and groups who are learning a language subsequent to learning their first one as young children” (p. 2). The language in question is called a second language (from here L2), in contrast to a first language (from here L1). A child’s first language is what we also call a mother tongue and is basically the first language a child learns; the parents’ role in this is the reason behind its second name. The second language does not only refer to the second language the child learns, but to any language that is learnt after the child has acquired a first language (Saville-Troike, 2006, p. 2). Saville-Troike (2006) makes a differentiation between informal and formal L2 learning, where informal learning happens in naturalistic contexts and formal learning is what we term classroom learning (p. 2). The method for learning explored in this study will be of relevance to both of these categories, as watching films and video clips is something that happens both in and outside of instruction in schools. Saville-Troike (2006) asks some thought provoking questions about SLA, and though they are not easy to answer satisfactorily they can be of value to teachers and learners. Two of the three questions asked are (1) ‘What exactly does the L2 learner come to know?’ (2) ‘How does the learner acquire this knowledge?’ (Saville-Troike, 2006, p. 2). These questions are the basis of many language research studies, and in this study I will explore the role of subtitles and video-clips in second language learning.

Writing in SLA

According to Brisk (2011), 'research has shown correlation between native language writing ability and performance in writing in English as a second language, even when the native language uses a different script' (p. 42-43). This correlation was not measured for in this study, but the suggestion of a connection of the two is interesting since writing performance is a factor that can influence the results of this study. English and Norwegian use the same script, but the idea of Norwegian as a predictor of English could mean that these two school subjects should consider this link as a tool for development in both languages. However, there is also another aspect to this relationship. Brisk (2011) also says that L1 can influence L2 products ‘especially grammatical structures and spelling, resulting in non-native like knowledge’ (p. 43). L1 influence can consequently contribute negatively to L2 production when users utilise their knowledge of their own L1 language where it does not belong.

According to Brisk (2011) it has been found that '[a]mong students who are learning English in an environment where their L1 is widely used, the influence of L1 may be greater' (p. 43). Naturally, Norwegian is the language that dominates in Norwegian schools, and the influence of it must therefore be considered to be large.

Grammar and vocabulary

We know that learners need vocabulary and grammar to produce language, but also to understand, and both of these aspects of language are therefore fundamental. Research has shown that the grammatical constructions we use when we speak vary a lot from those used when we write (Nation, 2011, p. 450). Nation (2011) writes that this is because complexity in speech is often largely casual, because of the amount of coordinated, noun and adverbial phrases, while written language is largely phrasal with a large number of complex nouns (p. 450). In Nation's (2011) interpretation, this means that written language (and formal language in particular) would not be adequate to the learning of spoken language, since the grammatical models would be different. Therefore, Nation concludes that teaching should focus on clausal rather than phrasal constructions (2011, p. 450). If this is a valid point, then film clips could be helpful in this context, because they often contain largely informal spoken language.

One of the issues in language teaching has always been the priority of certain areas, particularly of grammar and vocabulary. While there has always been a large focus on grammar, there 'has been a rise in the perceived importance of vocabulary as the main component in language proficiency, and by implication, in teaching' (Ur, 2011, p. 507). When vocabulary is seen as most important it is often in relation to communication as the essential aspect of language. Thus vocabulary plays a large role, since having knowledge of a vast amount of different words will presumably make it easier to communicate one's thoughts and meanings. For the purpose of this study vocabulary and grammar skills will be two variables in the investigation of language learning through film clips, based on the result of simple skill tests.

'Research exploring the variables that might contribute to L2 listening achievement is just emerging' writes Vandergift (2011), and refers to research by Mecartty (2000, as quoted in Vandergift, 2011, p. 460). In Mecartty's study vocabulary knowledge explained ca. 14 % of L2 listening ability, thereby appearing as a significant factor. By contrast, grammatical

knowledge failed in the same respect and was not a predictor. Vandergift (2011) speculates that this has to do with the two different categories' function, where vocabulary as content words is more significant than grammar as function words (p. 460). Pica (2011) states that 'grammatical progress has been shown to lag behind progress in listening, reading and oral communication skills' and if grammar skills lag behind in respect to other skills this could explain why grammar is not always a predictor (p. 262). Pica (2011) also writes that 'the consistent success of content and language integration has been revealed at the content and skill level' (p. 262). Content and function must thus be significant categories here, but this study does not measure any development in function/grammar skills; it is just measured as an initial factor. It could be claimed that vocabulary is in a sense measured both initially and later, as vocabulary size and comprehension of words. These measurements are not comparable since different methods and targets are used, but there could be a connection between them.

Nation and Webb (2011) claim that when one encounters a text, '[m]ost of the vocabulary in the text will be accounted for by the first 2,000 words of English' (p. 632). This would suggest that you need to know 2,000 words to understand most English texts (that are not targeted at a particularly advanced audience), which is not a very high number. In this study a level test has been used to establish vocabulary size, but as upper secondary students I suspect that they will know more than 2,000 words. This number might be relevant for simple understanding, but it will be interesting to see how vocabulary skills influence the results in this study. The tasks here can present a bigger challenge than the understanding of an ordinary text, since the target words are separated from the context. The context can in some cases make it easier to remember the meaning of the target word. Without knowing the content of these 2,000 words, it can be suspected that many of the target words fall outside of these.

Nation and Webb (2011) also list five conditions for vocabulary learning (p. 635-636). The first is motivation, which is considered an essential component for all types of learning. To watch a film can be a motivational activity in itself, because it will for many students act as variation from their ordinary school days of reading and listening to the teacher. Something that can become a snare for many teachers is that students may view a film mainly as almost a recreational activity, not as a means to learn. Unfortunately, this is a view is often shared by teachers. This can result in no other use of the clip and this can help to explain the students' attitudes. Some students may flinch when they hear there is learning involved in such a

setting, because they think that means more work, but a combination of learning and pleasure can be a motivation for some. Repetition is the second condition, and this implies several meetings with new words, perhaps also in different circumstances. If a film clip is on a specific topic, then some words can be mentioned several times and help learning because of this frequency. Nation and Webb's third condition is what they call the four strands - four important concepts of acquisition. Here input and output, and learning and development are central in interaction with each other. Thoughtful processing is the fourth condition, and reminds us that the learner should be aware of the learning process and consciously use strategies to pick up, learn and use new vocabulary. The fifth and last condition is called meaningful relationship, indicating the importance of creating relationships between words. To relate words to each other could facilitate the process of using them and understanding their meaning in different contexts. As a closing point here, we may note that Bird and Williams (2002) found in their study that bimodal input (of sound and text) 'can be attended to and used to bolster both the implicit and explicit aspects of vocabulary learning' (p. 18). This study will be referred to further in the chapter on previous research in the field of subtitles below.

Learning strategies in SLA

Jin and Cortazzi (2011) write about memory as a crucial factor in L2 learning, but, according to them, it is not necessarily connected to vocabulary (p. 569). By memory these authors mean explicit memory strategies or "memorisation" of words. From a cognitive perspective, it is standard to make a distinction between long term memory and (short-term)/ working memory. This type of memory is called long term memory, which is the permanent store of words and language concepts. We can call it a resource from where we take language units, which we use in the production of language. We use our working memory to manipulate the input we then store in our long term memory database. According to Perani (2005) the working memory (WM) system is a cognitive device, which we use to fulfil the requirements of daily life (p. 211). Perani (2005) writes that '[t]he phonological component of WM is considered to play a crucial role in language acquisition' (p. 211). The component of working memory called the phonological loop has been found to be a candidate for the language acquisition device, as phonological memory has shown to be important to both existing and developing vocabulary knowledge (Perani, 2005, pp. 211-212). Baddley (2000) has been central in the exploration of the function of the working memory and its main categories, where the phonological loop, the visuospatial sketchpad and the episodic buffer are central

terms (p. 417-418). The phonological loop is the most established component of this model of the working memory, and is relevant to this study as the component that helps store auditory information. According to Baddley (2000) it is a temporary store where auditory input quickly disappears if it is not rehearsed and end up in what we can assume is the long term memory of our mental lexicon (p. 419). Memorisation of vocabulary as a method for learning has 'suffered negative connotations of rote-learning from criticisms of grammar-translation' (Jin & Cortazzi, 2011, p. 569). The standard method in Norwegian schools has been to learn new vocabulary via texts, but there is still a lot of focus on the glossary at the back of the book and on the importance of memorizing these words, rather than focusing on the words and their meaning in the context. What Jin and Cortazzi (2011) suggests, based on other research, is that memory strategies are used, often spontaneously, and that they therefore are relevant. The difference could lie in the reasons behind the strategies, spontaneous or imposed. Video clips do not usually come with glossaries, and might thus inspire the use of memory strategies.

Chun (2011) writes about Computer-Assisted Language Learning; it is an area that is not directly related to the method used here, but can be seen in relation to subtitling and learning. According to Chun, research has shown 'that visual learners are guided by graphic information, but learners with (sic) low spatial ability are *not* helped by visual glosses of unknown L2 words (2011, p. 667, original emphasis). Also, Chun (2011) states that these low spatial-ability learners in meeting with both verbal and visual information that is to be processed do not learn vocabulary better (p. 667). If we transfer this to film watching and subtitling it seems as if the added subtitles will not help the weakest students to learn vocabulary, because they have to process different types of information. This issue is also addressed below. In relation to cartoons as Bahrani and Soltani (2011) concludes: 'Because of the type of modified language input which is embedded in cartoons, irrespective of the learners' proficiency level particularly low level language learners, they can respond instantly to what is being shown' (p. 21).

2.2 The role of input in SLA

Language heard or seen is called input. This is an important aspect of language learning, thus also of second language learning. According to Verspoor, Lowie and De Bot (2009), language acquisition cannot happen without input. If one already knows a language then proficiency can decline when it is not used, and even a first language can be forgotten (Verspoor et al.,

2009, p. 71). Hence as input of English is essential to start of English learning and developing it, it is also of major importance for maintaining proficiency. According to Vanpatten (2009) SLA is a slow process with sometimes incomplete results, because to make sense of a sentence “does not mean that all formal aspects contained in the utterance are fodder for acquisition” (p. 49). Even though input is very important in SLA, it is generally considered to be difficult for the ordinary L2 learner to acquire language solely through input, without output, interaction and instruction. Research suggests that for intermediate learners of English as a SLA, input can be beneficiary alone. Data from Verspoor and Winitz (1997, as quoted in Verspoor et al., 2009, p. 62) suggests that through just listening to English intermediate learners could improve their English in several areas, of vocabulary, grammar and reading skills. This is because our language changes constantly through interaction with the world around us, and the input we receive from this is part of a dynamic relationship with our language system (Verspoor et al., 2009, p. 71).

As Veerspoor et al. (2009) put it “first or second language development is an iterative process, which means that the present state of the learning system is the result of all previous steps or iterations” (p. 71). This is important for the SLA teacher, because the pupils will have been through different steps in the process of learning for example English, and will therefore be a group with a diverse collection of learning systems. For this study this is relevant because the participants were exposed to the same audio input within each group, but the study will show that there was a variation within the results of comprehension and word learning. Here the participants learning system at this time in their learning process will be different, and is a factor that plays an important role in their meeting with the language in the film clip. Learners who are at different levels will not process the same amount of what is provided by the input. An incipient learner could miss a lot of what is being said, while a more proficient learner will be able to process much more of what is said in the input (Verspoor et al. 2009, p. 73). Also, in addition to different initial conditions, input can be processed differently at different times (Verspoor et al. 2009, p. 74). This could be relevant for participants who are familiar with the series the film clip is taken from or any who have heard a word or idiom from it before, but who is now able to identify it from a context.

Krashen (2009) claims that input is more helpful than output in the learning situation. ‘Output fails as a predictor of second language competence when compared to reading; more speaking or writing does not result in more language or literacy development, but more reading does’

(Krashen, 2009, p. 85). If this is the case, then reading should be a prioritised method for language learning. For the learners, reading can get tiresome if there is not much variation and reading subtitles can be an alternative. However, as Verspoor et al. (2009) writes ‘the language system does not simply ‘take in’ input, but it interacts with it, adapts it, and reorganizes itself in the process’ (p. 77). Perhaps to use of the input in the production of output is a part of this process, and justifies output as important here too.

2.3 Subtitling of audio visual material and SLA

Subtitling is most known for its purpose of translating the oral dialogue from films or television programmes into text. This often involves viewers who have a L1 that is not the same as the language in the spoken dialogue and are given subtitles so that they can understand the action on the screen. This type of L2 subtitling is what we normally find in English speaking series on Norwegian television, where Norwegian subtitles are provided for the viewer. In addition to this, L1 subtitles are also sometimes available, though rarely on Norwegian television. Originally, L1 subtitles or captioning technology was devised for hearing impaired persons to aid their understanding when listening was not an option (Chiquito, 1995, p. 215). Today, it is also used by people who prefer English subtitles in addition to spoken English dialogue, instead of subtitles in their non-English mother tongue. In an instructional context it is used by teachers who want to take a step away from native language subtitles to challenge students as they become more proficient. It can also be the topic of research, as in this study, where people want to learn more about language learning and methods. Danan (2004) implies that subtitles can be a tool for teachers and an aid for learners:

Audiovisual material enhanced with captions or interlingual subtitles is a particularly powerful pedagogical tool which can help improve the listening comprehension skills of second-language learners. Captioning facilitates language learning by helping students visualize what they hear, especially if the input is not too far beyond their linguistic ability. Subtitling can also increase language comprehension and leads to additional cognitive benefits, such as greater depth of processing (p. 67)

This suggests that to see the orthography of the word aids learning of the word from the audio material, which justifies the view of subtitles as a useful tool in SLA.

Benefits of subtitles in SLA

Vanderplank (1988) was early in looking at the use of subtitling in language learning. In his study Vanderplank (1988) wanted to know if there were any benefits for language learning in watching programmes with subtitles (p. 272). Fifteen learners of English, who were between high-intermediate and post-proficiency level, watched BBC programmes with teletext subtitles and gave feedback on their experience with the experiment. The participants reported that the subtitles were beneficial to their language development and that ‘they were able to develop strategies and techniques for using sub-titles flexibly and according to need’ (p. 272). Vanderplank (1988) was therefore early to comment the use of subtitles in SLA, but suggests that subtitles are of more value to post-intermediate-level learners, providing comprehensible input, than for low-level learners (p. 272). The results also ‘indicated that sub-titles promote a low affective filter, encourage conscious language learning in “literate” learners, and, paradoxically, release spare language-processing capacity’ (Vanderplank, 1988, p. 272). That subtitles release capacity for processing and consequently learning in this way is an interesting point, as it implies a facilitation of acquisition.

In a later study Vanderplank (1990) again explores the benefits and limitations of subtitles as a tool in language learning. Vanderplank (1990) writes that what he found in his own study (from 1988, as presented above), and what others have found, suggests that the use of subtitles is not a distraction; ‘the double modal input appears to enhance comprehension better than simple script or sound’ (1990, p. 223). In his follow-up study Vanderplank (1990) reports that participants who paid attention to what they watched and used ‘note-taking or other aids to retention’ were helped by this and produced English of a higher level and improved their command of the language in social interaction (p. 226). For other participants who watched the programmes more passively, their recollection of terminology and details was limited and reproduction was not as good (Vanderplank, 1990, p. 226-227). Admittedly, this study is not new any longer, but the results are not necessarily outdated. Danan (2004) seems to be of the same opinion as she writes that ‘learners often need to be trained to develop active viewing strategies for an efficient use of captioned and subtitles material’ (p. 67). This is in accordance with Vanderplank’s (1990) findings on the relevance of actively engaging with the material.

Single modality vs. bimodal input

Bird and Williams (2002) name many studies that have measured the effects of same-language subtitling by showing video segments with or without subtitles, tested comprehension on plot and meaning of words and, in some studies, memory of exact phrasing has also been tested (p. 1). This study has not explored the last factor, but aims toward expanding the knowledge we have on subtitling by researching its effect in Norwegian. In other respects, this study follows the formula Bird and Williams refers to. Bird and Williams (2002) also write that they are not surprised by the results of these kinds of studies, where it is difficult to say if any improvements of comprehension could be due to reading subtitles or listening to spoken language (p. 2).

Bird and Williams' (2002) study consisted of two experiments that 'examined the effect of single modality (sound or text) and bimodal (sound and text) presentation on word learning' (p. 1). Their study found that 'simultaneous text presentation can aid novel word learning under certain condition, both as assessed by explicit and implicit memory tests' (Bird & Williams, 2002, p. 1). The tests in this study do not only look at novel words, but we can be quite certain that not all of the target words or phrases were in the participants' vocabulary already. According to Bird and Williams (2002), other research has suggested that students who have seen the written version of a spoken word will when presented with this word be better able to hear it (p. 2). The results from their study said that bimodal presentation was beneficial for implicit memory when new phonological forms needed to be encoded; Bird and Williams (2002) write that this suggested 'that the effect of text is limited to cases where the phonological form of the input cannot be reliably established on the basis of sound alone' (p. 17). In dealing with familiar words, textual support was not needed as the low error rates showed that sound input alone was sufficient (Bird & Williams, 2002, p. 17). The study does not make it clear if there was any interaction and influence between the two modalities in the processing of the words, but Bird and Williams (2002) expected that the modalities would compensate for any deficiencies in the other modality. In the case of explicit memory Bird and Williams (2002) found that bimodal input aided the performance of explicit memory in a recognition task (p. 18). In subtitling literature it has been conscious recall that typically is tested, but Bird and Williams (2002) removed the semantic context from the subtitles in their experiments and thereby 'show more clearly that providing subjects with text and sound versions of known and unknown words can facilitate recognition memory relative to sound alone' (p. 18).

Video clips with subtitles –reading or listening?

In his study, Markham (1999) found ‘positive evidence concerning the influence of second-language captions directly on second-language listening’ (p. 326). Markham (1999) does stress that since ‘the clear majority of students were Asian, generalizing the results of this research to other ESL populations should be done very cautiously’ (p. 322). Vandergift (2011) writes that ‘[t]he consensus of research conducted on the usefulness of captions for listening comprehension is that L2 captions facilitate comprehension’, but that it is still not clear how well it fares for long-term effects of learning, in the case of listening improvement and vocabulary learning (p. 462). Vandergift (2011) questions the use of subtitles, since it could inhibit development of listening skills; ‘students will not *learn* how to listen if they read to understand L2 aural texts’ (p. 462, original emphasis). Therefore the decision of subtitle use in the classroom could be made according to the purpose of the film showing. If the focus is on listening comprehension, Vandergift (2011) suggests that subtitles will hinder this. Bird and Williams (2002) also comments this issue by writing that ‘it remains unclear whether subtitles are leading to better or worse listening comprehension’ (p. 2), but that the results from their study showed ‘that the bimodal condition created no apparent interference with auditory processing and learning (p. 18). Also it ‘lead to improved implicit learning of novel word forms when the experiment allowed targets to be recycled three times prior to test’ (Bird & Williams, 2002, p. 18). The memory experiment of this study is not as extensive as this, but it measures memory of spoken words within groups of different subtitle conditions.

Views on subtitling and captioning

According to Gunderson, Odo and D’Silva (2011) same language subtitling (SLS) has been a success in rural India (p. 480). It has been helpful in providing literacy, also for L2 literacy (L2L), and has promoted reading through a simple and effective method for language development. This report is from an area that it not as literal as urban areas, but it is still interesting to note as language learning is an important issue in all levels of society. Bird and Williams’ (2002) experiment showed that orthographic information could help the learning of spoken words, also with long terms effect for implicit and explicit memory. This experiment had participants with both western and eastern background, and shows that this method could be useful for people of several nationalities. Bird and Williams (2002) suggest using the same-language subtitles that most DVDs now provide, even if they are originally intended for the hearing impaired, for second language learning (p. 19).

According to Danan (2004) the hostility that subtitling sometimes meets for example in the United States comes about because of accusations ‘of encouraging viewers to rely on the written text, taking attention away from the actual spoken language, and even fostering a form of laziness bordering on cheating’ (p. 67). Yet, as Danan (2004) writes: ‘many Europeans claim to have learned English from their regular exposure to subtitled American films and television programs’ (p. 67). While ‘the reading of subtitles tends to be an automatic behaviour’ this does not mean that it prevents any processing of the soundtrack (Danan, 2004, p. 72). There can be limitations to captioning; Guillory (1998) found that captions are no longer useful when the material is too advanced, and the captions cannot compensate for difficult vocabulary and fast speech (as quoted in Danan, 2004, p. 71). Danan (2004) still claims that captions can be valuable for comprehension, word recognition and vocabulary building (p. 69).

Subtitles and accents

As mentioned in the introduction, an article by Mitterer and McQueen (2009) inspired the creation of this study. In their study they ‘investigated whether subtitles, which provide lexical information, support perceptual learning about foreign speech’ (p. 1). Mitterer and McQueen (2009) used video clips with Scottish and Australian regional accents of English, and here differs from this study which used a more standard American accent. The first language of the participants was Dutch and they were also older and more advanced, as university students who used English reading materials (Mitterer & McQueen, 2009, p. 4). The testing involved showing two groups a video clip with either Scottish or Australian speech, and then have the participants repeat audio fragments in both accent variants (Mitterer and McQueen, 2009, p. 5). Mitterer and McQueen (2009) found that the participants who saw the clip with English subtitles performed better than those who saw it with Dutch subtitles and this applied to both new and old items. Old items were taken from the video clips, while new items were taken from the same source as the clip, but from unwatched material. They found that the ‘benefit due to the English subtitles thus appears to reflect generalization of learning across the lexicon’ (Mitterer & McQueen, 2009, p. 3). In summary, they ‘tested whether audiovisual exposure allows listeners to adapt to an unfamiliar foreign accent [... and] whether subtitles can influence this process’ and found that ‘this kind of adaptation is possible and that subtitles which match the foreign spoken language help adaptation while subtitles in the listener’s native language hinder adaptation’ (Mitterer & McQueen, 2009, p. 3). The results showed that English subtitles led to better performances: ‘Perceptual adaptation was enhanced

by subtitles that were in the same language as the accented speech. Adaption effects, and their enhancement by English subtitles, were found for old and new items' (Mitterer & McQueen, 2009, p. 3).

The teacher

Vanderplank (1990) found in his study that the teacher is important in the watching of subtitled programmes; the study reports that 'most subjects required some explicit purpose, some degree of instrumental motivation in order to attend to the language of the programme being shown and take out words and phrases' (p. 227). Simple exposure to language seems to be sufficient for learners of the language; the process needs to be more conscious and therefore requires more of the learner. This suggests that the teacher should provide the learners with some basic instruction and motivation to create more interest for the task.

2.4 Visual material: cartoons/animations

In this study we used an animated cartoon called *Family Guy*. Animation is 'the art of making inanimate objects appear to move' (Encyclopædia Britannica). Bahrani and Soltani (2011) claim that cartoons have been effective in increasing the motivation of language learners, as input of authentic audiovisual language (p. 19). Bahrani and Soltani (2011) write that 'cartoons are good sources of modified language input which although may require less cognitive processing for comprehension, they may include new aspects of the language for low level language learners who have not acquired them yet' (p. 19). Chun's (2011) view on subtitles was that it did not help low spatial learners, and the use of a cartoon with subtitles can perhaps modify this claim. According to Bahrani and Soltani (2011) cartoons should not be used as a time-filler by for example a teacher, but needs to be relevant when used in teaching and the use can be justified by other effects, as the belief that 'the usage of cartoons can reduce boredom and decrease academic stress, anxiety and disruptive behaviour' (p. 19). As well as creating an atmosphere that causes motivation, exposure to cartoons has been claimed to enhance memory when connecting new and old materials through analogy (Bahrani & Soltani, 2011, p. 19-20). Bahrani and Soltani (2011) encourage the use of cartoons in lessons because they provide variation for the brain. The explanation for this is that the spoken word engages the left side of the brain, and as this is what is most typically used in instruction the learners will feel the dullness of it after a while (Bahrani & Soltani, 2011, p. 20). A cartoon, on the other hand, is a type of visual information and is processed by the right

side of the brain and is therefore a tool for variety and effective in keeping learners less bored (Bahrani & Soltani, 2011, p. 20). For a study that has used a cartoon as material for testing effects of subtitles see Bianchi and Ciabattini (2008).

2.5 Priming

Priming is a simple automatic effect where a connection is seen between different stimuli and is quite often exploited in experiments studying the mental lexicon. One of the tests in this experiment is a lexical decision task where one of the objects is to look for signs of priming. Evidence of priming has been found in many other studies, even among very young individuals. Mani and Plunkett (2010) investigated in their study ‘whether infants can implicitly name visually fixated images’ after other studies had found evidence of similar priming in adults (p. 908). Even though the lexicon of the infant (18-month-olds) is smaller than that of an adult, they could ‘test whether a prime image influences an infant’s response to an image with a label that begins with the same phoneme as the prime label’ (Mani & Plunkett, 2010, p. 909). The phonological relationship between the unheard prime label and the heard target label was the only attested relationship between targets and primes in the primed trials (Mani & Plunkett, 2010, p. 909). According to the results of this study, when the prime and target words were phonetically and semantically unrelated the infants did not show target recognition, but did so when there was a semantic relation between the prime and target (Mani & Plunkett, 2010, p. 911). Thus, this priming experiment showed positive effects of priming on very young participants.

Fischler’s (1977) study had a lexical decision task where the subjects were asked to decide if a stimulus was a word. These were presented as a pair of letter strings, and the task was to say if both of them were words. The materials consisted of pairs that were associated with each other, words that were semantically similar and unrelated control pairs, and ‘[b]oth associates and semantically related pairs were responded to more quickly than were the corresponding control pairs’ (Fischler, 1977, p. 335). Fischler (1977) concludes that ‘the encoding of a word can be facilitated by the prior processing of a semantically related word’ (p. 335). Fischler’s study is indeed many years old now, as Shelton and Martin (1992) pointed out already 15 years later. According to Shelton and Martin (1992) it was afterwards discovered that ‘much of the priming in a lexical decision task comes about due to subject strategies rather than to

automatic processes' and this was because 'subjects notice that some primes and targets are related and use this information to aid their word-nonword decisions' (p. 1191).

The most important result that Shelton and Martin (1992) drew from their own study was that automatic priming did not happen for words that had a semantic relation, but no association, while it did come about for associated words. Therefore Shelton and Martin (1992) conclude that 'words that are very similar in meaning or sharing many features will not show automatic semantic priming if they are not also associated' (p. 1204). They argue that 'this associative priming might result from connections between lexical phonological or orthographic representation that have developed on the basis of co-occurrence frequency' instead of as a result of connections at a meaning level (Shelton & Martin, 1992, p. 1207). However, McRae and Boisvert (1998) showed in their experiments a 'convincing demonstration of a "pure semantic" priming effect' (as quoted in Ferrand & New, 2004). McRae and Boisvert's (1998) results indicated that reliable priming depended on methodological considerations and 'demonstrated that featural similarity is an important organizing principle of semantic memory' (p. 570).

Ferrand and New (2004) created an experiment 'designed to examine semantic memory by focusing on two types of "semantic" relatedness, associative relatedness and semantic similarity' and found that 'on average, the size of the priming effect was larger for semantic pairs [...] than for associative pairs' (p. 37). Ferrand and New (2004) write that these results were consistent with what the spreading-activation theory of Collins and Loftus (1975) predicts in that '[t]he automatic semantic priming would result from spreading activation between semantic representations' while 'the automatic associative priming would result from spreading activation between (non-semantic) lexical representations' (Ferrand & New, 2004, p. 38). Ferrand and New (2004) suggest that semantic and associative relatedness can provide priming under automatic conditions, though separately from each other (p. 38).

3.0 Method

During the beginning of the project, and occasionally throughout the process of collecting and using the data for the project, I have been collaborating with another master student who is writing a similar thesis. The difference between our projects is that we have used groups of participants from two different levels at upper secondary, with expected different levels of proficiency. Of my two supervisors we had one in common, which made this collaboration highly appropriate. Because of this I will in this chapter refer to things that I have done and to things that we in collaboration have done together, hence the use of I/we to distinguish this.

3.1 Research questions and the quantitative method

Research questions are important, as helpful guiding tools when collecting and analysing data (Sunderland, 2010, p. 9). For this project, I wanted to see if different subtitles influenced understanding of a short video with English speaking audio material. Since I and my fellow master student wanted to keep the opportunity of a potential comparison of our results across projects, we decided to co-operate on the questions asked and the way of collecting data. This comparison will not be treated here, but most probably in a later paper. The main issue has been to look for a hypothetical influence of subtitles on comprehension and language learning. As already mentioned, the idea of researching the use of subtitles was inspired by the earlier study by Mitterer and McQueen (2009). This study was interesting, and inspired in us a wish to look for similar results with Norwegian, but we would use a more standardized American English rather the dialects used in their study. Additionally, younger participants were used, which further differentiates our studies. Mitterer and McQueen (2009) used methods that focused more on the oral aspect of English (admittedly because they used English accents in their material), while our studies have used written test materials. The study of Mitterer and McQueen (2009) and this study are not directly comparable in terms of result, but this relation will be treated in the general discussion. Another source for research questions is one's own findings, which according to Sunderland (2010) is a more controversial method (p. 12). Since I am looking for differences in the results I could find these in areas I did not expect to.

When one develops tests the researcher always has to decide on a method, and whether one will use primarily quantitative or qualitative method. Angouri (2010) deals with this issue, and also presents a combination of the two as an alternative (p. 29-42). Though some of the

questions in the background questionnaire asked the participants to define an answer themselves, suggesting a qualitative approach, the method used in this study is the quantitative method. As Rasinger (2010) comments, quantitative makes us think of ‘quantity’ as opposed to ‘quality’, where the first refers to ‘much’ and the latter to ‘good’ (p. 50). This would be a misunderstanding; quantitative methods can be used satisfyingly in research, as both qualitative and quantitative methods have their areas of operation. Qualitative research is more concerned with the question of *why*, while the focus of quantitative method is more numerical and concerned with the questions *how much* and *how many* (Rasinger, 2010, p. 52). The latter questions are the most important for this study, as we want to establish how, or if, the second language learner use subtitles as a learning tool. Questions of why would also be interesting, but not appropriate for what we seek to find out in this study. As Rasinger writes “[q]uantitative data can be analysed using statistical methods, that is, particular mathematics tools which allows us to work with numerical data” (p. 52). For this study, as we wanted to use many participants, such a tool was preferred, because it presents a relatively easy way of analysing data. Thus *Microsoft Excel* and the statistical programme *R* were used for handling and analysing data. The scope of this project demanded an amount of participants to get a conclusive result, and quantitative method presented itself as the most logical choice. Admittedly, qualitative method was not to a large degree considered already from the beginning of this process, which could be because our focus and take on this project was clear from an early stage. This type of research is deductive, where one sets out with a hypothesis based on already known theory (Rasinger, 2010, p. 52). In our case we set out with assumptions that were not based on our own study of theory, but as suggested by our supervisor. We used a type of experimental design, which means that we could manipulate the variables and thereby approve or disprove the hypothesis, which cannot be done in other types of designs, for example in observational designs (Rasinger, 2010, p. 59).

3.2 Participants.

Participants from three classes at an upper secondary school took part in this study; the classes had an average of 30 students each and the participants were 16 years old. Among these students I have used the results from a total of 65 participants and of these 34 were girls and 31 were boys. For my study I have used students from VG1, the lowest level in the Norwegian upper secondary school system, within the study programme specialisation in general studies (Norwegian: *studiespesialisering*). At this level English is an obligatory subject and students have not yet chosen their specialisation. In the two studies by me and the

other master student, the participants of this study were the younger group. This choice of participants was based on the assumption that there would be a difference in the English level between these two groups.

The three groups I used were from three classes respectively, and the classes each had a different teacher. One of the teachers became our contact person and she created the lists of the participants. These lists gave each of the participants a number, and as researchers we had no access to their names. Not all of the students in each class were part of the study I present here, either because they did not attend both parts of the testing, or they had to be excluded for the results because they did not fit within the target group. Based on the information from the background questionnaire, participants were excluded from the study if they had or once had problems concerning hearing, other language learning problems or diagnoses that could influence language learning (ADHD, autism etc.). Participants who had another mother tongue than Norwegian were also excluded, because of the design of the study where Norwegian and English were used in the subtitles, as the participants' L1 and L2. In the end, group 1 had 25 participants, group 2 had 16 participants and group 3 had 24 participants. It was not easy to predict that this was the number of participants I would end up with. The classes were large enough that I considered that at least 20 persons per group could be used for the study, but these numbers did vary across the groups with both more and less participants. As mentioned in the introduction the participants were not selected as individuals and put into groups, but were selected for this study as the classes they receive educational instruction in. The initial same size of the three classes made them well suitable to be used in the same form, as three separate groups for this round of testing. Rasinger (2010) presents the experimental group and the control group as part of an experimental design, described above (p. 59). I therefore had two experimental groups that were undergoing the stimulus, which in this study was to watch the film clip with subtitles (Norwegian or English), and one control group who saw it without subtitles and were not affected by the stimulus.

The participants in this project were under the age of majority, and their participation therefore had to be approved by their parents. A parent's permit was therefore sent home with the pupils (appendix A). The permit included a short description on what we were going to do with this project, without revealing the precise details, for example it said that we would show a video clip and did not specify from which film or series this clip came from. The important function of the document was to inform the parents about the participants' anonymity, that we

would never see their names, but give each participant a code. These codes would only be identifiable via a list, kept by their respective teachers, and this list would be deleted once the information was gathered. The permit also specified that the student could withdraw from the testing at any time, so that the student too could choose to not be part of the project. This consent letter also contained information on when the testing would happen, a short description of the information we would gather from the students and contact information for me and my colleague. It also specified that the results from the study would not influence the grades of the students and that the teacher would not have access to the results.

3.3 Materials

When this study was planned, we expected to find differences in the results between the groups with different subtitles, so the question was about which strategies we would use to look for these. We were looking for differences without knowing where these would appear, so we decided to use more than one test. We ended up using four tests or questionnaires, which will be presented below. Even though we expected there to be a difference between the groups, we did not want to rule out other potential variables. This is why we collected info about various aspects of the linguistic background and English proficiency level of the participants. We created a comprehension questionnaire, to check the participants' comprehension of the happenings in the film, as this was relevant for the language testing. For the main part of the testing, we used both a word definition task and a lexical decision task.

Before the tests were created, we had to find a suitable video clip to show and base our tests on. We aimed at finding a cartoon to use, because the genre would hopefully motivate the participants to pay attention to the clip. A cartoon was considered to be a better option than for example a documentary, which would perhaps not capture all parts of the audience. Documentaries are entertaining for many, but this was a group of young people who have a lot of different interests, and since the clip and this study was not related to their grade we sought to encourage an interest to watch in other ways. In this thought process, Family Guy, created by Seth MacFarlane, was chosen as the source of a clip. The series has episodes that are approximately twenty minutes long, and this makes one episode suitable as a relatively short video clip that still contained enough language to create our tests from. First, we had to choose an episode that would be appropriate to show among pupils at upper secondary level. Many of the episodes from Family Guy contain inappropriate language and motifs, so we had

to choose one that was relatively inoffensive. Our choice was an episode from season 7, called ‘Movin’ out (Brian’s song)’. We then transcribed the episode in three versions: spoken text, English subtitles, and Norwegian subtitles. This gave us the opportunity to study the script and the subtitles side by side, to seek phrases and words that it would be relevant for us to use in our studies.

Background questionnaire

According to Rasinger (2010), the questionnaire is a frequently used method within linguistic research, as it can be used when asking for opinions on language (p. 60). It is important that the questionnaire is a planned, well thought through piece of data gathering.

The background questionnaire (appendix B) used for this study was created from an already existing version used in language research at the Department of Modern Foreign Languages, a well-established and successful questionnaire. It was adapted to our use by adding questions that asked about English habits and proficiency, questions on language and mother tongue, and about other factors that could potentially affect language learning. Initially, this questionnaire was not considered the most important piece of our testing, but as important to mapping the proficiency of the participants, and to find possible exclusions among these. This view was later modified. As the results of the tests related to the video clip were analysed, it was apparent that the data collected from the background questionnaire would be central in explaining the results from the following tests, as will be seen below. In the questionnaire most of the questions asked for the extent of something, with possible answers on a scale from “every day” to “never”. Another task was to place their skill level within different areas of English on a scale, with four levels from basic to fluently. This was also done for other second languages, though only for the general level of language and not for separate areas. Some questions were more open, because they asked the participants to name for example countries they had visited, or the types of computer games they played. As the theories of Verspoor et al. indicates, the participants would have different learning systems because of their previous experience with language learning and dynamic process with input, as this process is iterative. Thus the background questionnaire is an important tool for establishing their position in relation to English learning, by mapping their use of English and encounters with the language. The process the participants have been through to come to their level as we met them in this study is, of course, too complicated and perhaps impossible to map, but through asking about their habits with English one gets a good idea of how this dynamic process influence them and their learning system.

As part of the background testing, we also had the participants do two types of computer based tests that were available on the internet. These measured the participants' level of proficiency, more specifically vocabulary size and grammar proficiency, and these were factors in the analysis. The first of these tests was a vocabulary test that was to measure the size of the participants' vocabulary (<http://dynamo.dictionary.com/placement/level>). This test asked for the appropriate definitions for ten different words, where there were four alternatives to choose from. If one answered with the wrong definition the size of the calculated vocabulary would shrink, as it would increase when the correct alternative was chosen. The website offers four different tests divided by school level and we chose to use the version created for middle school. In America, middle school students are around 11 to 13 years old, thus significantly younger than the participants of this study. As English was our participants' second language, middle school was an appropriate level to choose here. The two higher levels would have proven to be too difficult for the participants, which could have resulted in guessing or unfair scores and thereby giving us faulty measurements. The participants did two rounds of this test, because we wanted to have a more valid result that evened out lucky guesses. To have the two rounds counted on the same vocabulary score demanded a registering of an account; this is something that would have taken time and demanded more from our participants, therefore they did two separate rounds and I calculated the average score of these for each participant. The second test was a grammar test with fifty exercises, where the end score was given as the amount of correct answers (http://www.cambridge.org/other_files/Flash_apps/inuse/EssGramTest/EssGramIndex.htm). This tested essential grammar, which was preferred for this study in contrast to a more advanced version. Each exercise contained a sentence with a section missing, this was either a word or a combination of words, where there was from two to four alternatives to select from. The possible answers were very similar in orthography and/or meaning, but only one alternative was grammatically correct in the target sentence. This test was only taken once and the top score was fifty points.

Comprehension questionnaire

For the first round of testing we made a short comprehension questionnaire (appendix C), which the participants were to receive and answer immediately after watching the film. This test would check if the participants understood important parts of the plot. This was one of two comprehension tests, but the only one that tested comprehension directly after the input in the video clip had been watched, thus at an earlier stage in the processing of the input

received in the clip. The questionnaire had 18 exercises to be answered. These were formulated as unfinished sentences that had four alternative continuations, which the participant was to choose from. The questions were fairly short and uncomplicated. One example is the sentence that began with ‘Peter and Jillian go to see’ where the correct continuation in this case was ‘Disney on Ice’.

Word definition task

The main research question in this study concerns whether Norwegian or English subtitles make it easier for the students to learn English and thus we needed to look at both short term and long term effects. To look for any long term effects we created a word definition task (appendix D) where the participants would choose between four definitions of a word or a phrase; this made it the second test concerned with comprehension. The task was presented in written form, and the participants would circle their chosen answer: a, b, c or d. This was a multiple choice task, thus the participants did not define the stimuli in their own words, but chose among the alternatives. One task was “To hire”, and the alternatives were “a. To get rid of an employee, b. To tie a metallic rod, c. To employ someone, d. To advance to a higher position”, the answer obviously being c. Among the alternatives there was one correct answer, two similar but incorrect replies and one that was distinctively incorrect. There were a total of thirty exercises in this test and eight of these were idioms. The words and phrases were all taken from the transcribed text of the film clip, and were chosen because they were interesting words; some are more frequently used in American English, while others were a bit rarer. We wanted to know if the participants would understand their meaning by watching the film clip and learn it well enough to remember it weeks later, and also if there were any differences across the groups. All the words and phrases were looked up in the COCA (see below) to evaluate their frequency in language use, which showed that our selection represented different areas on the scale.

Lexical decision task

We decided to create a lexical decision task (appendix E) to look for long term recollection of words they had heard in the audio material of the video clip. In the creation of this task we had to make a choice of which words to include, and we needed a selection of words that were uttered in the clip and a selection of words that were not. Some of the words that were not mentioned were still in this episode through visual priming, and we wanted to see if the participants would mark these as being heard in the episode. To find a range of words that

could be placed on different sides on a scale of frequency, we used a corpus. The one we chose was ‘The Corpus of Contemporary American English (COCA)’, as the variety of English used in the film clip is mainly American English (with the exception of the character Stewie, who has a British accent). The COCA was not used for analytical purposes, but as a tool in determining if our selection of words were representative of different levels of frequency.

For a subtest within the lexical decision task we also chose a selection of words that were not mentioned in the episodes, but were English synonyms of English words used in the audio material. The Norwegian words in the Norwegian subtitle were originally translations of the English audio material, and we back translated these words into English by using a different word than in the original audio material. This was done to see if we would find traces of priming through Norwegian, since the participant with the Norwegian subtitles could favour back translations of the words. For practical purposes these words were in the lexical decision task given to all the three groups, but in the analysis of the results this type of priming was only looked for in the data material from the Norwegian subtitle group.

3.4 Procedure

During the testing period for this study we visited the groups twice and these visits were four weeks apart in time. I collaborated with my colleague in the testing period too, since the groups used in our separate projects often had their classes at the same time, making it impossible to visit several groups when the time is limited. It was preferable to visit all the groups within approximately the same time period and the visits were made at the time when the participants usually had instruction in English. Both rounds of testing spanned for two days in order to get through all the groups for both projects. For the first visit, we had to do testing in three classes at the same time, so an external research assistant was brought in from outside of the project and given specific instruction to make sure that his procedure would be the same as ours. The second visit required a shorter time slot and further assistance was not necessary. At no point in the testing period were the participants told what we as researchers were looking for in our study.

In the first part of the testing period we mapped the participants’ language background with the background questionnaire and the subsequent grammar and vocabulary test. Most of the

testing in this period was done on paper, except the grammar and vocabulary tests that were answered on a computer. The students had been asked to bring their own school computer, as Norwegian students at upper secondary level get support for this, and used it to answer the tests that required it. They were instructed to answer the background questionnaire and when finished do the computer based tests. Steps with instructions and web addresses were also written on the whiteboard. To make sure that no one would cheat and write up a higher score on the vocabulary and grammar tests, the participants had to raise their hand each time they finished one test. In this way the person representing the researchers or the teacher could write it down for them, to control its correctness. We did not give a set time slot for the different tests, but continued by starting the video clip when everyone had finished the two steps above. The participants were not told to look for anything in particular in the clip, as they were only told that a video clip would now be played. As mentioned, this clip lasted approximately 20 minutes. Immediately after it ended we handed out the comprehension questionnaire, not giving the participants time to speak among themselves on what they had seen. Everyone remained quiet until the last person had finished this questionnaire. At the end of this round of testing we told the participants that we would come back over Christmas (this was in December), but not any details on what would happen then.

The second round of testing took less time, as there were only tests on paper and no more video clips to be watched. Students from the group that had not been present in the previous round of testing were excused from class. In this round the participants were given both tests at once, which were printed together in the same booklet with the word definition task first and the lexical decision task as the last page. Before we handed out the booklet instructions on how to answer were given and as the last was handed in we thanked for their cooperation in helping us with our projects. All in all, the testing period was enabled by the collaboration between the researcher and the cooperation of the teachers and participants.

3.5 Validity.

For this experiment and study to be of significance to research, the results have to be valid, in other words reliable. According to Postholm (2010), a normal criterion for reliable results is that they can be reproduced and repeated (p. 169). This is what Nunan (1992) calls external validity (p. 15). The study can be reproduced, but the results could vary in other participant groups, as these results are taken from participants from a specific language/school

background. Nunan (1992) also writes of internal validity, which concerns factors that can influence the outcome of a study directly (p. 15). That the participants are young is a factor of the study that the researcher must take into consideration; communication and instruction must happen in a certain way. For the researchers in this study, this was not a problem. Since we are teachers we are used to dealing with pupils or students, and we know how to give clear instructions to avoid confusion in the group. The instructions were therefore to the point, both those given orally and on the tests. We also wrote the steps on the whiteboard, in case someone was not paying attention or forgot. On our first visit with the groups we said that we were going to do some test and watch a video clip, and on the second visit that they were going to answer some questions relating to what we had done last time. The purpose with the experiment and the focus on subtitles was never mentioned to the participants. To the best of our knowledge, they were therefore unaware of our intentions. We have to assume that the replies we have collected are sincere and consequently that the data is reliable.

4.0 Results

All data that could not be used for this experiment has been excluded before the analysis, for reasons described above. All other collected responses were coded into data sheets in Microsoft Excel, the results from the test as “y” or “n” for correct and incorrect. As mentioned above, mathematics tools are utilisable in the process of analysing quantitative data. Excel was used to organise the data and facilitate the further analysis of it. The data was inspected in the program *R*, in a generalized linear mixed model fit by Laplace approximation. Each of the tests has had their results analysed using *R*, and some of the tests had additional aspects that were investigated and are reported below. The models were created by using the factors from the initial background testing of language skills/habits and vocabulary/grammar results, where the variables that were significant to the results were kept in the final models. I used likelihood tests (function `anova()`) to determine which model to keep from each of my analyses. The tests results were explored separately, and the ratio between the different groups varied from test to test, and the models will therefore have different variables as predictors of the results.

4.1 Background

Table 1|Results from initial vocabulary and grammar tests

	English subtitles	Norwegian subtitles	Control group
Vocabulary	11698,72	14842,91	11015,65
Grammar	42,68	44,88	44,33

In the first round of the testing, the participants each did a vocabulary test twice and a grammar test once, as explained above. The average scores of these tests are presented here by group. We immediately see that there are similarities between the groups. As table 1 shows, the vocabulary results for the English subtitle group and the control group are very close, while the Norwegian subtitle group stands out with a higher average score. All the groups have an average score that is well above 2000, which was the number Nation and Webb (2011) gave as required to understand most of an “ordinary” English text. The grammar score are also similar between two of the groups, while the English subtitle group scores

slightly lower. The maximum score on the grammar test was 50, thus the groups scored as was expected for their level: quite high, but not flawless.

4.2 Comprehension questionnaire

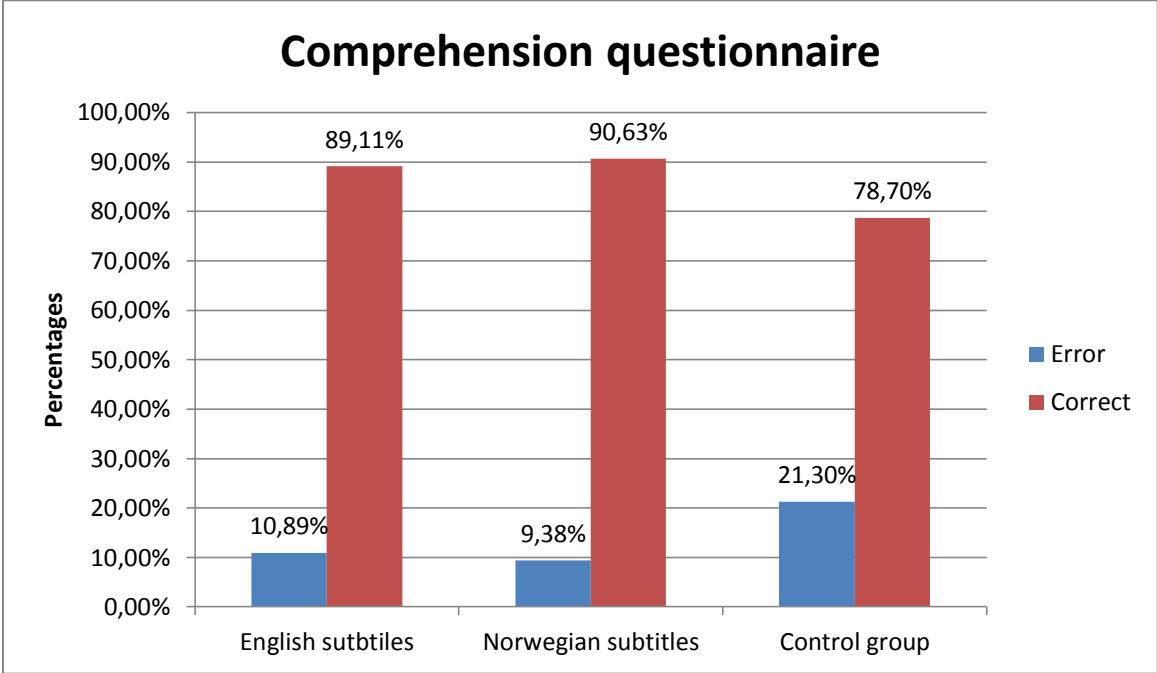


Figure 1| Average scores on comprehension questionnaire by group

The results from the comprehension test are visible in the bars of figure 1. All of the three groups score quite high, but the control group stands out with a higher error average. The comprehension test was considered by its creators to be fairly easy, and high averages of correct answers were expected. However, it seems from the result that the two groups who watched the clip with subtitles scored highest. This indicates that subtitles helped participants' performance on the comprehension test.

Table 2| Results from analysis of comprehension test

	Estimate	SE	z-value	Pr(> z)
(Intercept)	-12.74289	3.50561	3.635	0.000278 ***
GroupEng	0.78831	0.29695	2.655	0.007938 **
GroupNor	0.75125	0.36342	2.067	0.038722 *
log(voc)	1.33337	0.38792	3.437	0.000588 ***
listen_eng	0.65233	0.20861	3.127	0.001766 **
eng_cg	0.18479	0.09895	1.868	0.061831 (*)

(*) < 0.1; *<0.05; **<0.01; ***<0.001

From table 2 we can see that the expectation based on figure 1 can be justified, the subtitles did help the participants in their performance on this test. As the comprehension questionnaire demanded an understanding of the plot of the video clip, this would suggest that the written subtitles aided this type of comprehension. The control group did not have a text to aid them and subsequently had to rely on the audio material, which resulted in less comprehension. The table also shows that the English subtitle was a more significant predictor for this group than the Norwegian subtitle was for the other group. This is interesting, since the groups that watched with Norwegian subtitles performed better on average. The most significant predictor for the comprehension test is the vocabulary score from the background testing, participants with a higher vocabulary score performed better on the comprehension test. This could explain why Norwegian subtitles are less significant than English subtitles in the analysis, since the first group had higher vocabulary scores. This will also have an effect on the analysis of the word definition task below. The participants' self-evaluated skills in listening to English does also show up as a predictor for this test. To have a good understanding of spoken language does therefore predict a better performance on this test. There is also a small hit on the significance of computer games. This variable measured how often the participants played English speaking computer games, and this factor seems to have marginally predicted performance here; this means that those who played computer games more frequently benefited from this in some way.

Table 3| Comprehension test as predictor for Word Definition test

	Estimate	SE	z-value	Pr(> z)
(Intercept)	0.1031	0.3898	0.264	0.791
accuracycompCy	0.5327	0.2144	2.484	0.013*
accuracyLDy	0.1233	0.1736	0.710	0.478

compC= comprehension questionnaire, LD= lexical decision task. *<0.05; **<0.01; ***<0.001

While exploring the data I wanted to see if any of the tests correlated with the results of the other tests. When examining the data from the word definition task I found that the comprehension test showed a significant positive correlation. As table 3 shows, answering correct in the word definition task is connected to answering correctly in the word definition task. This could be expected, since both of them test comprehension, of the action and words/phrases respectively. This does not have to mean that one test functions as a predictive variable to the other test. There is technically a correlation here that suggests a connection between the tests, where performance in the comprehension questionnaire correlated with performance in the word definition task, but with no correlation with the lexical decision task. However, I view this as a correlation, not a predictor. Hence this is not included in the original model for the word definition task presented in table 4 below, since the relation between the tests has more to do with the variables behind them than their corresponding results.

4.3 Word definition task

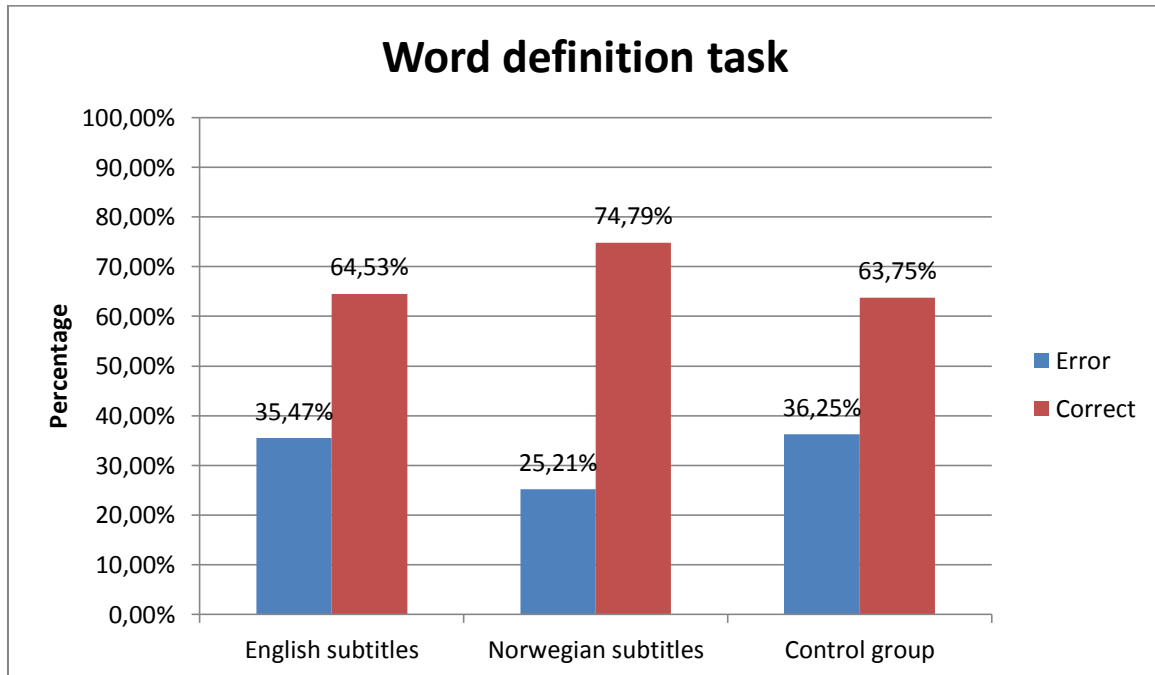


Figure 2 | Average results on word definition task by group

The result of the word definition task appears to show partly what was initially presumed of the benefits of Norwegian subtitles. In figure 2 we see a higher average score for the Norwegian subtitles, while both of the other two groups have on average more incorrect replies. From just looking at this figure, it would seem that the Norwegian subtitles were beneficial in the word definition task. What should be remembered is that the same group scored higher than the others on the vocabulary test, as seen above, which could mean that the groups were different on the outset. On the other hand, the English subtitles seem to not have any benefits for the participants, as their scores are very close to the control group who watched without subtitles. The explanation could be that the participants simply did not pay attention to the subtitles, or they did read them, but it did not give them an advantage over watching the clip without it.

Table 4| Results from analysis of word definition task

	Estimate	SE	z-value	Pr(> z)
(Intercept)	-10.12281	1.49018	-6.793	1.10e-11 ***
GroupEng	-0.12572	0.17645	-0.713	0.4762
GroupNor	0.28114	0.20975	1.340	0.1801
log(frequencyWD)	-0.13184	0.09394	-1.403	0.1605
log(vocabulary)	1.08642	0.14594	7.444	9.74e-14 ***
GroupEng:log(frequencyWD)	-0.01348	0.05028	-0.268	0.7886
GroupNor:log(frequencyWD)	0.11875	0.05847	2.031	0.0422 *

Variable log(vocabulary) is a combination of two variables: vocabulary and self-evaluated written English skills, as these showed a correlation. * <0.05 ; ** <0.01 ; *** <0.001

As we can see from table 4, group does not show up as a significant predictor of response accuracy in this model, and was consequently not a predictor of the results. The only reason that group appears in the model here is because I have looked at the interaction between group and frequency of words. Frequency signified the frequency per million of the words in the test, based on COCA. Neither frequency is significant here as a predictor for the results overall, across groups; the only hit in this respect is for the Norwegian subtitle group. It seems that the group that watched the video clip with Norwegian subtitles performed better when the frequency of the word was higher, while that did not help the group with English subtitles. The most striking aspect of this model is the significance of the vocabulary factor. As noted, this variable consists of the data from the vocabulary test and the participants' self-evaluation of their skills in writing English, because these correlated in the model. This means that participants that scored high on the vocabulary test and thus as a large vocabulary, and considered themselves to have good skills of written English (probably within the advanced or fluent category) had more accurate replies than others.

Figure 2 implied that Norwegian subtitles would be significant since that group scored higher on the word definition task. We also knew from the background questionnaire that the same group had a higher average score on the vocabulary test, and from this model it seems that this was important to the results. High vocabulary scores must therefore have been a favouring factor in the other groups as well as in the Norwegian subtitle group, and thus it predicted the results, while the subtitles had no such influence on the results.

4.4 Lexical decision task

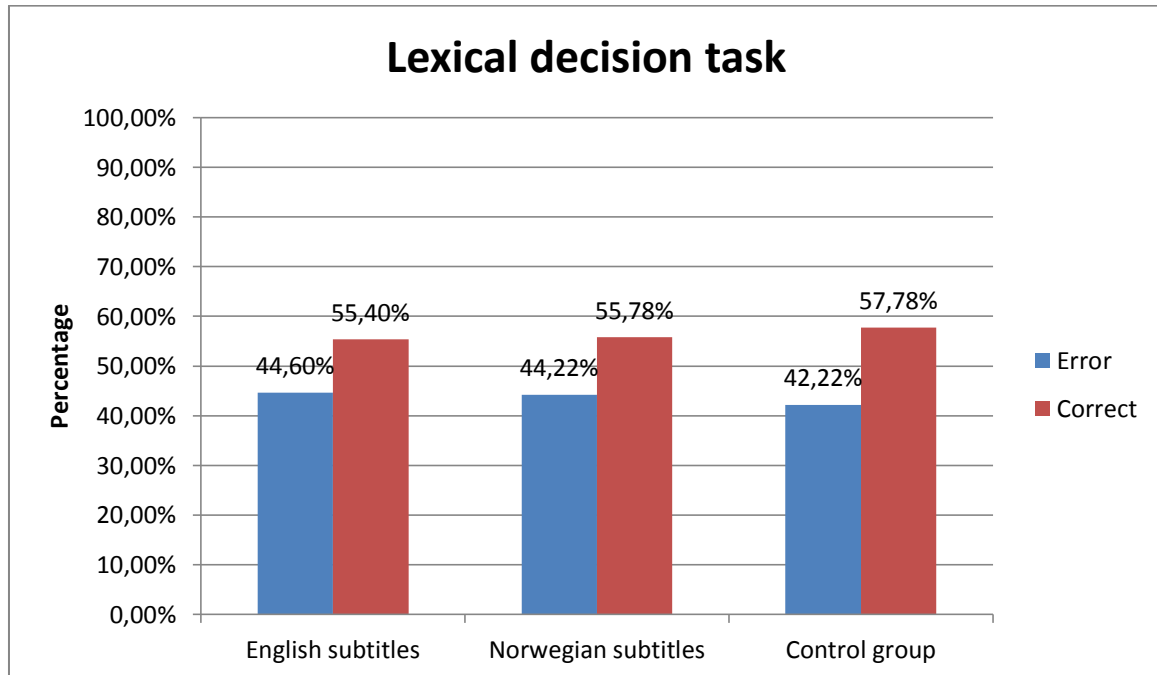


Figure 3 | Results from lexical decision task by group

Figure 3 here indicates that group is not a predictor for the lexical decision task, since the groups average scores are almost equal. The control group has the lowest score, but is still only 2 % below the Norwegian subtitle group, which scores marginally highest on this test.

Table 5| Results from analysis of Lexical Decision task

	Estimate	SE	z-value	Pr(> z)
(Intercept)	0.95448	0.32403	2.946	0.00322 **
cartoon	-0.09498	0.03544	-2.680	0.00736 **
write_engT	-0.11107	0.04723	-2.352	0.01868 *

Write_engT is a factor that indicates how often the participants write English text. *<0.05; **<0.01

The analysis of the lexical decision task, as seen in table 5, shows that fewer factors emerged as significant predictors of the results in this test, than in the others. Cartoon appears to be a significant factor, which means that the participants who watched cartoons more often scored better on this test. This indicates that exposure to authentic material is of relevance here. To write English text often is the other predictor here, as slightly less predictive than cartoons.

Table 6| Results from priming experiment in Lexical Decision task

	Estimate	SE	z-value	Pr(> z)
(Intercept)	0.2558	0.3126	0.818	0.413
primeLDp	1.1085	0.6904	1.606	0.108
primeLDvp	-1.3350	1.0033	-1.331	0.183

LD= lexical decision, p=prime, vp=visual prime.

One part of the experiment was to look for effects of priming in the lexical decision task. This concerned only the results from the Norwegian subtitle group, thus the results from this experiment seen in table 6 is based on data from this group only. Among the words in the lexical decision task there were, as mentioned, some words that were back-translated to English from the Norwegian subtitle. These words were marked as ‘p’ and from table 6 we see that this variable was not significant to the result of this test for the target group. Visual prime was a marker for words that appeared in the video clip through a visual representation, but were not explicitly said in the clip. This priming factor was not significant either, neither for this group nor for the three groups overall. The only thing we can read from the analysis of the visual prime marker is that it influenced the result of the participants negatively. One example is the stimulus ‘dog’, which was often chosen by participants. One of the characters in the video clip was a dog and this was never mentioned in the clip, but the visual representation would have confused some participants to this respect, or as we could say; primed a faulty reply

5.0 Discussion

The process that the participants have been through in this study can be summed up in the words of Bianchi and Ciabattini (2008):

When students watch a film in a foreign language and text aids are displayed, three channels compete in catching the students' attention and in favouring (or hampering) comprehension and learning: one auditory channel, and two visual channels (one verbal and one non-verbal). In this scenario, several different variables are at play, including the following: semantic match between the verbal channels (audio and text) and the non-verbal channel (images); type of text aid (captions, subtitles, no text aid); student level of proficiency; and type of task (content, vocabulary, or language-in-use comprehension or acquisition) (p. 86).

The different types of input through sound, text and picture make it a complex situation, where different variables can help or harm the processing of these. What this study has investigated is which variables predict positive results on the created tests. The comprehension questionnaire targeted content, the word definition task targeted vocabulary and comprehension, while the lexical decision task targeted acquisition/memory. It was initially assumed that group would be a significant factor in this, because of the design of the study where the groups watched different subtitle versions of the same clip. This assumption did not confirm itself in the results of each test; it was rather other variables that stood out and these were largely connected with students' level of English proficiency. In addition, the expected priming situation was not realised in the results. The predictors for the result do offer a nuanced view on the testing situation, as will be seen in the following that directly discusses the results of the different tests.

5.1 Comprehension questionnaire

In the results of the comprehension questionnaire group appears as a predictor of performance. This was an experimental condition, where the three groups watched with either English subtitles, Norwegian subtitles or with no subtitles. Significantly it is the two groups

that watched with subtitles that performed best here, while the control group scored lower on comprehension. This would imply that the textual information provided in the subtitles aided the understanding of the plot. We had expected high scores on this questionnaire, since the questions were not considered to be very difficult. Also, as Bahrani and Soltani (2011) writes the cartoon contains modified language that can be instantly responded to irrespective of proficiency level and this was an immediate test of comprehension after viewing the clip. The control group who did not have access to this verbal visual channel was inhibited by this. Initially the gap between the subtitle groups and the control group does not appear to be very big, from what we visualise from figure 1. When we examine the numbers the gap is approximately 10% and as the questionnaire had 18 stimuli; the average difference is almost two (1.8) correct answers in advantage Norwegian subtitles. Accordingly, L1 influence did not contribute negatively to L2 comprehension. It is interesting is that the Norwegian subtitle is a less significant predictor than the English subtitle, even though the results seen in figure 1 show that the group with Norwegian subtitles scored marginally higher on the comprehension questionnaire. As already mentioned, this could be seen in connection to the vocabulary scores seen in table 1. The Norwegian subtitle group scored significantly better on the vocabulary test, and as this test score shows up as the most significant predictor in the results it is reasonable to assume that the subtitles were not as important as a predictor as the vocabulary score was, in the case of the Norwegian subtitle group. This would also mean that even though the results of the English subtitle group were also predicted by the vocabulary score, the relation between the score and the predictiveness of the subtitles vary from that of the Norwegian subtitle group. Vandergift (2011) saw the significance of vocabulary, then as opposed to grammar, as an indicator for the value of content words over function words for the learner's listening ability. The experiment of this study uses more than the audio channel, but we see similar results. Grammar knowledge does not show up as significant, while vocabulary knowledge predicts performance overall. Vocabulary knowledge and its relation to the subtitle variable will also be discussed below, as it was an even more significant predictor of the results from the word definition task.

It is of interest that vocabulary knowledge was found to predict listening ability, as this skills shows up as a predictor in itself for the comprehension questionnaire. In the background questionnaire the participants were asked to evaluate their English proficiency within different areas, where listening was one area. It seems then that good listening skills predicted test scores. Here the phonological loop is an important device for language acquisition, as it is

the component of the working memory that functions as a temporary store of auditory input. Listening skills appears as more significant than for example the Norwegian subtitles and slightly more than the English subtitles as well. What we can assume from this is that the participants benefitted more from the input from the audio channel, the spoken dialogue of the clip, than the textual information from the verbal visual channel of the subtitles. Listening ability could thus be a significant predictor for the control group that did not have any textual information to rely on. If so, then learners of English with good listening abilities should be able to comprehend video both with and without subtitles to a relative degree.

From the results, computer games also show up as a predictor of performance. This means that to play English speaking computer games often was beneficial for this task, though marginally and less so than the other predictors above. This is still an interesting result to comment on, since computer games as we see them today are of a relatively modern and technological era, and provides an opportunity of language interaction as was not available in earlier years. Even if the genre of computer games goes back some decades, the games that we see today are far more advanced than in its beginning. According to Ang and Zaphiri (2008) the player of a computer game interacts with the game as it gives the player instruction of how to play and what to do, which the player adapts to and a player might need to interact with other players. Here English is often the language that most players share and use in interaction with international players. Thus input will be available mostly from the game, but also from other players. In considering the interaction process and use of the input in the light of theory from Verspoor et.al (2009) we can assume that this language situation can help improve the player's English. This then could have helped the participants' understanding of the content of the video clip and influenced the comprehension results for this part of the testing. As it was not a predictor of the later comprehension test, the word definition task, we can consider it as a predictor for general comprehension, though not as comprehension at word level.

5.2 Word definition task

As we saw in figure 2, the Norwegian subtitle initially seemed to be a significant predictor for the word definition task, as this group scored significantly higher than the other two groups, with a differentiation of approximately 10%. As this task had 30 stimuli, this gives a difference of three correct answers. However, the results from the analysis as presented in

table 4 tells a different story, as group does not show up as a significant factor. The group variables would not have been present in the final model at all if frequency of words had appeared as significant when viewed in relation to groups. As this did happen, the table shows that group is no longer a significant factor. This is a surprising result as we had expected the subtitles to predict the results. As the participants' watched a short video clip which was the only input, and the tests were based on this input, it was expected that the subtitles would be of importance. The different versions of the clip with Norwegian subtitles, English subtitles or no subtitles was the only variable that was controlled by us from the start, and thus it was assumed that there would be traces of this difference within the results.

As it turns out, subtitles do not predict performance for long term effects on comprehension. The high score on the vocabulary test must therefore predict performance for the Norwegian subtitle group here as in the comprehension questionnaire. This is not surprising as vocabulary and the mental lexicon are big factors in language learning. In this test, the vocabulary score outshines the Norwegian subtitle to the degree that it does not show up as significant at all. This is relatively more dramatic, especially as we consider the hypothesis and the assumptions based only on figure 2. In light of the significance of vocabulary and the score of the Norwegian subtitle group in the initial background testing, we can confirm that the vocabulary score here predicts performance, at the expense of the Norwegian subtitle. This is not the only interesting finding. From table 4 we find that the English subtitle was not significant in predicting performance and from figure 2 we can see from the average results that the score between the English subtitle group and the control group differed by less than 1%, though in advantage of the first group. As mentioned, this would imply that the group did not benefit from the English subtitles and the suggested explanations were that they either did or did not use them, but they did not aid long term comprehension. It is tempting to suggest that the participants of this group did not find any help in the subtitles and consequently did not use them, but then subtitle was such a significant factor for the comprehension questionnaire in contrast to the no subtitles of the control group. From this we can extract that the group did read the English subtitles, but that they were of no benefit for long term comprehension. Also, Chun (2011) has written that subtitles with unknown words do not aid low spatial ability learners in understanding them. The word definition task requires more of the participant's comprehension at a word level, and these learners could have managed sufficiently in the comprehension questionnaire, which also was close in time to the viewing, but not on this test.

In the model presented in table 4 the variable 'vocabulary' is a combination of the vocabulary test score from the background questionnaire and the self-evaluated proficiency of written English. This variable is the most significant one both within the task and across all tests. According to Ur (2011) vocabulary knowledge has increasingly become the main component of proficiency, and this is confirmed in these results. The reason for this could be vocabulary's nature as content words. The comprehension questionnaire and the word definition task are both concerned with comprehension, of either plot or words and phrases. Grammar as function words is perhaps not of such crucial importance for comprehension, as function is more relevant to small details. It has been seen to lag behind other skills, which suggests that it is less important to proficiency. Written English skill is also a significant predictor of performance and as it correlated with vocabulary score. A large vocabulary would lead to a better written language that is varied and more advanced, thus intermediate learners would have better written skills as seen in relation to a vocabulary size. This we can assume from the triangular interaction between phonological representation, orthography and meaning. Krashen (2009) claims that reading, input, is more helpful than writing, output, in the learning situation, but here reading skills is not a predictor, while writing is. We can see this in the light of what Verspoor et al (2009) write about interaction with the input. Thus it can be the ability to work and interact with the input that shows up as a predictor through writing in this task.

Group does show up as a significant factor in relation to another factor, namely frequency. The frequency variable is based on per million frequencies of each stimulus, as found in the COCA. As it was only a predictor for one group, it seems that the Norwegian subtitle group performed better on words that have a high frequency in contemporary language use. This we can see in relation to the group's average vocabulary score, in that the frequent words were perhaps part of their vocabulary span already. As a contrast, the English group did not benefit from higher frequency, as would have been to a higher degree expected from them since they on average scored lower on the word definition task and on the vocabulary test. Logically, a smaller vocabulary should have predicted more correct answers for high frequency words, but this is not visible here. Then vocabulary size must also be significant for knowledge of frequent words.

Interestingly, upon analysis of the relation between the tests there was found a correlation between the comprehension questionnaire and the word definition task, which indicates that

there was a connection between these. Thus the participants' score on the one test correlated with the score on the other. Still, based on the analysis of the tests we can say that subtitles and vocabulary score predicted performance on the comprehension questionnaire, while subtitles were passed as predictors of the word definition task in favour of vocabulary scores.

5.3 Lexical decision task

The analysis of the lexical decision task showed cartoon and frequent writing of English text as predictors of performance on this test. There were no clear assumptions for the general results of this task, though it was perhaps expected that the English subtitles would aid memory of words. As mentioned, Bird and Williams (2002) found in their study that simultaneous text presentation showed positive results for both explicit and implicit memory. As Danan (2004) claimed, a visualisation of a spoken word could aid learning, while Nation and Webb (2011) stress the importance of repetition. It might be a bit bold to suggest that a word is repeated because it appears both in the audio and in the subtitle, but it is available in both channels. According to Jin and Cortazzi (2011) language learners often use memory strategies spontaneously. Thus even though the participants were not instructed beforehand to remember certain words, it would be expected that some type of memory mechanism would still be activated. That subtitles was not a predictor suggests that the participants did not benefit from having both a visual and an auditory version of the word in this memory experiment.

That the participants who watched cartoons more often benefited from it in this task, might be because of the type of input. As they were used to watching cartoons, they had perhaps unconsciously developed strategies in dealing with cartoons that created an advantage for them here. Frequent viewing of the target series, *Family Guy*, was not a significant predictor, so it was not the language of that particular series that helped participants choose the correct words. What I know of the advantages of cartoons does not help explain this predictor in this context, other than as a type of familiar input and that they through cartoons are exposed to authentic language material. To write English more often was also a predictor and must also be of importance because of interaction with input, as discussed with the variable of written English skill for the word definition task above.

As can be seen in table 6, visual priming was not a significant predictor for the Norwegian subtitle group, and the analysis showed no significance for the whole data set either. Though not significantly, the values show a negative influence of visual primes on the results in table 6. Many participants marked *dog* and *bedroom*, two targets that were very visible in the clip, but never explicitly named in the audio (nor in subtitles). The confusion of a heard or seen word could mean that the participants were not conscious of where they had picked it up, from the audio, text or visual material. Any significance for this was not found, thus we must assume that on average the participants were not influenced by this visual priming.

The main prime experiment of the lexical decision task was priming for Norwegian-English back translations, which was based on the Norwegian subtitle and the audio script, and analysed with data from the Norwegian group only. In Fischler's (1997) study the word pairs were presented as pairs, while in this experiment the target back translated word was only given four weeks after the original English word had been heard in the audio material and the Norwegian word read in the Norwegian subtitle. In reality we here have three words, where the expectation was that the English target word would be confused with the English word from the audio material, because it had the same meaning. This expectation was not realised. We might have expected an automatic process of priming, but the words were semantically similar and Shelton and Martin (1992) claim that priming is strategic for semantically related words. This requires the subject to notice a relationship, and we can from this deduce that the participants did not notice this relation and did not adopt a strategy to solve this task.

5.4 General discussion

Mitterer and McQueen (2009) found that native language subtitles harmed foreign speech perception, while it in this study did not harm perception, though it neither significantly aided it. It could be that their experiment demanded more of the participants, as it used accents and the English subtitles made the spoken language more familiar to its participants. Also, the participants were university students and perhaps did strain themselves more than our participants who were pupils and used to being tested.

This study aimed to find traces of the influence of subtitling on language learning and we therefore sought to keep the influence from our own role at a minimum. Motivation is viewed as an important factor for most types of learning, and the use of video can be a variation in the

learning situation that students can appreciate. The important thing for the instructor is to have a purpose with showing the clip and as Vanderplank (1990) suggests, to give instructions on how to approach it so that the students can use strategies for retention of language input. Perhaps the participants would have had more use of the subtitles if they had made notes after viewing the clip, which would have contributed to the iterative acquisition process. Then again, note taking can mean that the students miss something in the clip. Most higher education students have experienced taking notes in a lecture and when writing we have missed something from the lecturer's speech. If such strategies are to be used, it must be after consideration of the clip and the information the instructor wants the students to acquire from it.

An important factor to language learning seems to be that of the outset of the learners, the learner's level will be a decisive factor in their level of performance. The general consensus is that the incipient learner will not profit as much as the more proficient learner, when one considers the amount of input one is able to process (Verspoor et.al, 2009; Chiquito, 1995). Guillroy (1998) also pointed out that subtitles do not help if the language is too advanced. Research such as by Verspoor and Winitz (1997) has found that for intermediate learners, listening alone can help improve their English. This leaves out the incipient learners, but Danan (2004) suggests that subtitles can help these learners because they provide orthographic information in addition to the heard word. This is what Bird and Williams (2002) calls 'simultaneous text presentation', where they suggest that the different modalities could compensate for deficiencies in each other. Additionally, according to Bahrani and Soltani (2011) cartoons provide modified language input that can reach learners at low levels as well. We cannot say for sure which modality the participants used the most, but for the comprehension questionnaire it seems that both subtitle groups benefited from these for comprehension of plot, while the control group lagged a bit behind. This suggests that reading the subtitles aided simple comprehension. Also, research has shown that after seeing a written version a word the student will be better able to hear it (Bird and Williams, 2002). A possible method is to have students read beforehand a text on the topic that uses much of the same vocabulary, or with subtitles to show the clip first with and then without subtitles. This would, of course, depend on factors as the length of the clip.

What we did not find was any influence of the subtitles on the tests of round two, where we tested for long term effects, which is closer to what acquisition really is - a more stable

outcome. The consensus is that L2 captions facilitate comprehension, but as Vandergift (2011) writes the long-term effects are more uncertain and this study did not find any signs of long terms effects because of the subtitles. This could be why teachers mainly use video clips for informational purposes rather than specifically for language learning. However, as Danan (2004) writes, there are many who claim to have learned English through exposure to it via films and TV. Thus the results of this study must not be seen as a warning against the use of subtitled video, but the instructor must consider the level of the students, as we have seen that input and vocabulary size are important predictors of learning in this method.

We cannot say that one of the subtitle versions worked better than the other. For this project we only used data from participants with Norwegian as their mother tongue, but in the real classrooms there are increasingly more students of different language backgrounds. Their English acquisition should not suffer because their Norwegian is not proficient enough, and then English subtitles can be used. Gunderson et al. (2011) found that same language subtitling had been a success for literacy in India and Bird and Williams (2002) found that people of several nationalities benefited from it. Finally, video material provides variation for the student and helps engage more parts of the brain to make learning more stimulating.

6.0 Conclusion

In this study we aimed to get a better knowledge of the influence of subtitles on second language acquisition for Norwegian learners of English. We used a twenty minute video clip of a cartoon and found that subtitles aided comprehension of plot, which was tested immediately after the screening. English and Norwegian subtitles aided comprehension, which was found in comparison with the control group that watched with no subtitles. Vocabulary size, which was tested for initially, turned out to be a significant predictor of general comprehension. It predicted the results on both the comprehension questionnaire and the word definition task. It was expected that subtitles would also be a predictor of the results of the word definition task, a presumption that was not confirmed, as the outset of the participants proved to be a more important factor. The proficiency level of the participants, based on their skills in interaction with the input and their vocabulary range, was a determiner for the result. This suggests that subtitles do not affect long term language learning, and that the learner's proficiency level is a more significant factor for further learning.

The memory experiment of the lexical decision task showed similar results across groups. Since the subtitles were not a predictor this suggests that the participants did not benefit from seeing the textual version of the word in addition to hearing it, when they tried to remember if it had been uttered or not. Cartoon as a type of input and regular writing of English were predictor of performance, but do not explain the results other than as interaction with input and strategies adapted when watching cartoons. The prime experiment with back translation did not show signs of priming, and showed that the Norwegian subtitle group were not confused by the Norwegian subtitle, which initially was a translation of the English script.

Bibliography

- Ang, C. S., & Zaphiris, P. (2008). Computer Games and Language Learning. In T. T. Kidd & H. Song (Eds.), *Handbook of Research on Instructional Systems & Technology* (pp. 449-462). Hershey, PA: IGI Global.
- Angouri, J. (2010). Quantitative, Qualitative or Both? Combining Methods in Linguistic Research. In L. Litosseliti (ed.), *Research Methods in Linguistics*. (1.edition, pp. 29-45), London : Continuum.
- animation. (2013). In *Encyclopædia Britannica*. Retrieved 19 April, 2013 from <http://global.britannica.com/EBchecked/topic/25806/animation>
- Baddeley, A. (2000). The episodic buffer: a new component of working memory? *Trends in Cognitive Sciences*, 4(11), 417-423. doi: 10.1016/S1364-6613(00)01538-2
- Bahrani, T. & Soltani, R. (2011). The pedagogical values of cartoons. *Research on Humanities and Social Sciences*, 1(4), 19-22.
- Bianchi, F. & Ciabattini, T. (2008). Captions and Subtitles in EFL Learning: an investigative study in a comprehensive computer environment. In A. Baldry et al. (ed), *From Didactas to Ecolingua: an Ongoing Research Project on Translation and Corpus Linguistics*. (pp. 69-90). Trieste: Edizioni Università
- Bird, S. A. & Williams, J. N. (2002). The Effect of Bimodal Input on Implicit and Explicit Memory: An Investigation into the Benefits of Within-Language Subtitling. *Applied Psycholinguistics*, 23(4), 509-533. doi: 10.1017/S0142716402004022
- Brisk, M.E. (2011). Learning to Write in the Second Language: K-5. In E. Hinkel (ed.), *Handbook of Research in Second Language Teaching and Learning, Volume II*. (pp. 40-56), London: Routledge

- Chiquito, A.B. (1995). Metacognitive Learning Techniques in the User Interface : Advance Organizers and Captioning. *Computers and the Humanities*, 28(4-5), 211-223.
doi: 10.1007/BF01830268
- Chun, D. M. (2011). Computer-Assisted Language Learning. In E. Hinkel (ed.), *Handbook of Research in Second Language Teaching and Learning, Volume II*. (pp. 663-680), London: Routledge
- Collins, A. M., & Loftus, E. F. (1975). A spreading activation theory of semantic processing. *Psychological Review*, 62(6), 407-428. doi: 10.1037/0033-295X.82.6.407
- Danan, M. (2004). Captioning and Subtitling: Undervalued Language Learning Strategies. *Meta: Translators' Journal*, 49(1), 67-77. doi: 10.7202/009021ar
- Ferrand, L. & New, B. (2004). Semantic and Associative Priming in the Mental Lexicon. In P. Bonin (ed.), *Mental Lexicon: Some Words to Talk About Words*. (pp. 25-43). New York: Nova Science Publishers
- Fischler, I. (1977). Semantic facilitation without association in a lexical decision task. *Memory & Cognition*, 5(3), 335-339. doi: 10.3758/BF03197580
- Guillroy, H.G. (1998). The effects of keyword captions to authentic French video on learner comprehension. *Calico Journal*, 15(1/3), 89-108.
- Gunderson, L, Odo, D.M. & D'Silva, R. (2011). Second Language Literacy. In E. Hinkel (ed.), *Handbook of Research in Second Language Teaching and Learning, Volume II*. (pp. 472-487), London: Routledge
- Jin, L. & Cortazzi, M.(2011). Re-Evaluating Traditional Approaches to Second Language Teaching and Learning. In E. Hinkel (ed.), *Handbook of Research in Second Language Teaching and Learning, Volume II*. (pp. 558-575), London: Routledge

- Krashen, S. (2009). The Comprehension Hypothesis Extended. In T. Piske & M. Young-Scholten (ed.), *Input Matters in SLA*. (1. edition, pp. 81-94), Bristol: Multilingual Matters
- Mani, N. & Plunkett, K. (2010). In the Infant's Mins's Ear: Evidence for Implicit Naming in 18-Month-Olds. *Psychological Science*, *21*(7), 908-913.
doi: 10.1177/0956797610373371.
- Markham, P. (1999). Captioned Videotapes and Second-Language Listening Word Recognition. *Foreign Language Annals*, *32*(3), 321-328. doi: 10.1111/j.1944-9720.1999.tb01344.x
- McRae, K. & Boisvert, S. (1998). Automatic Semantic Similarity Priming. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *24*(3), 558-572.
- Mecartty, F. (2000). Lexical and grammatical knowledge in second language reading and listening comprehension by foreign language learners of Spanish. *Applied Language Learning*, *11* (2), 323-348.
- Mitterer, H. & McQueen, J.M. (2009). Foreign Subtitles Help but Native-Language Subtitles Harm Foreign Speech Perception. *PLoS ONE*, *4*(11), e7785, 1-5.
doi: 10.1371/journal.pone.0007785
- Nation, I.S.P. (2011). Second Language Speaking. In E. Hinkel (ed.), *Handbook of Research in Second Language Teaching and Learning, Volume II*. (pp. 444-454), London: Routledge
- Nation, I.S.P. & Webb, S. (2011). Content-Based Instruction and Vocabulary Learning. In E. Hinkel (ed.), *Handbook of Research in Second Language Teaching and Learning, Volume II*. (pp. 631-644), London: Routledge
- Nunan, D. (1992). *Research methods in language learning*. Cambridge: Cambridge University Press

- Perani, D. (2005). The neural basis of language talent in bilinguals. *TRENDS in Cognitive Sciences*, 9(5), 211-213. doi:10.1016/j.tics.2005.03.001
- Pica, T. (2011). Second Language Acquisition Research: Applied and Applicable Orientations to Practical Questions and Concerns. In E. Hinkel (ed.), *Handbook of Research in Second Language Teaching and Learning, Volume II*. (pp. 257-273), London: Routledge
- Postholm, M. B. (2010). *Kvalitativ metode. En innføring med fokus på fenomenologi, etnografi og kasusstudier* (2. edition). Oslo: Universitetsforlaget (Norwegian)
- Rasinger, S.M. (2010). Quantitative Methods: Concepts, Frameworks and Issues. In L. Litosseliti (ed.), *Research Methods in Linguistics*. (1. edition, pp. 49-67), London : Continuum.
- Saville-Troike, M. (2006). *Introducing Second Language Acquisition*. Cambridge: Cambridge University Press
- Shelton, J.R. & Martin, R.C. (1992). How Semantic Is Automatic Semantic Priming. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 18(6), 1191-1210.
- Sunderland, J. (2010). Research Questions in Linguistics. In L. Litosseliti (ed.), *Research Methods in Linguistics*. (1.edition, pp. 9-28), London : Continuum.
- Ur, P. (2011). Grammar Teaching: Research, Theory and Practice. In E. Hinkel (ed.), *Handbook of Research in Second Language Teaching and Learning, Volume II*. (pp. 507-522), London: Routledge
- Vandergift, L. (2011). Second Language Learning: Presage, Process, Product and Pedagogy. In E. Hinkel (ed.), *Handbook of Research in Second Language Teaching and Learning, Volume II*. (pp. 455-471), London: Routledge
- Vanderplank, R. (1988). The value of teletext sub-titles in language learning. *ELT Journal*, 42(4), 272-281. doi: 10.1093/elt/42.4.272

- Vanderplank, R. (1990). Paying attention to the words: Practical and theoretical problems in watching television programmes with uni-lingual (CEEFAX) sub-titles. *System*, 18(2), 221-234. doi: 10.1016/0346-251X(90)90056-B
- Vanpatten, B. (2009). Processing Matters in Input Enhancement. In T. Piske & M. Young-Scholten (ed.), *Input Matters in SLA*. (1. edition, pp. 47-61), Bristol: Multilingual Matters
- Verspoor, M., Lowie, W. & de Bot, K. (2009). Input and Second Language Development from a Dynamic Perspective. In T. Piske & M. Young-Scholten (ed.), *Input Matters in SLA*. (1. edition, pp. 62-80), Bristol: Multilingual Matters
- Verspoor, M. & Winitz, H. (1997) Assessment of the lexical-input approach for intermediate language learners. *IRAL*, 35 (1), 61-75.
- Viener, J. (Writer), & Tang, C. (Director). (30. September, 2007). Movin' out (Brian's Song) [Television series episode]. In K. Vallow (Producer), *Family Guy*. Los Angeles, CA: Fox Broadcasting Company

Appendices

Appendix A – Parent’s consent form

Samtykke til deltakelse i undersøkelse om andrespråkforståelse

Ansvarlig institusjon: NTNU

Studenter: Lisa Aurstad (lisaaurstad@live.no) og Ingrid Kvitnes (ingrid.kvitnes@gmail.com)

Veiledere: Mila Vulchanova og Juhani Järvikivi

Vi ønsker å gjennomføre en undersøkelse i elevens klasse der vi med utgangspunkt i engelsk som andrespråk vil se på elevenes kompetanse i og forståelse av det engelske språket.

Studien vil bestå av to deler og begge vil foregå i skoletiden. Del 1 vil foregå i desember (uke 50), og del 2 vil foregå i januar (uke 2). Del 1 innebærer at elevene skal være med på en kort kartlegging av sin språklige bakgrunn og kompetanse i engelsk, samt at de skal se et filmklipp. Del 2 inneholder språklige tester knyttet til filmklippet.

En kode knytter eleven til sine opplysninger gjennom en deltakerliste. Det er kun autorisert personell knyttet til prosjektet som har adgang til deltakerlisten og som kan finne tilbake til informasjonen. Det er kun læreren som har oversikt over hvilke navn som er knyttet til koden. Læreren vil ikke ha tilgang til elevens resultater i studien, og studien vil derfor ikke kunne ha innvirkning på elevens karakterer. Denne oversikten vil også bli slettet når studien er ferdig (15.5.2013). All informasjon vil bli anonymisert ved prosjektslutt, og alle opplysninger gitt i undersøkelsen vil bli behandlet konfidensielt. Det vil ikke være mulig å identifisere eleven i resultatene av studien når disse publiseres. Også skolen vil bli anonymisert.

Selv om eleven ikke skal delta i studien, vil eleven måtte være til stede i klasserommet under gjennomføringen av studien. Eleven får altså ikke fri til å gjøre andre ting mens undersøkelsen pågår. Deltakelse i undersøkelsen er frivilling, og eleven kan når som helst trekke seg fra undersøkelsen underveis uten å oppgi en spesiell grunn.

Vennligst fyll ut og lever denne samtykkeerklæringen til elevens engelsklærer.

Vi ber om at skjemaet leveres så raskt som mulig for at eleven skal kunne delta.

Ta kontakt med Lisa eller Ingrid for spørsmål.

Jeg samtykker til at _____(elevens navn) kan delta i undersøkelsen.

Trondheim, Dato: _____ Underskrift: _____

Appendix B – Background questionnaire

Bakgrunnsinformasjon for forskningsprosjekt om andrespråkforståelse

Tusen takk for at du har sagt ja til å delta i vårt forskningsprosjekt om andrespråkforståelse. I dette skjemaet ber vi om bakgrunnsinformasjon som er nødvendig for at resultatene fra undersøkelsen skal kunne brukes.

Alle opplysningene du gir her, vil senere bli behandlet uten direkte gjenkjennende opplysninger. En kode knytter deg til dine opplysninger gjennom en deltakerliste. Det er kun autorisert personell knyttet til prosjektet som har adgang til deltakerlisten og som kan finne tilbake til infoen. Del B, C og D av dette skjemaet vil bare oppbevares med koden. All informasjon vil bli anonymisert ved prosjektslutt. Det vil ikke være mulig å identifisere deg i resultatene av studien når disse publiseres.

Legg merke til at skjemaet har 6 sider.

Med takknemlig hilsen,

Lisa Aurstad / Ingrid Kvitnes

Studenter ved lektorutdanning med master i språk, NTNU

Del A: Personlig informasjon

Studieretning og trinn: _____

Fødselsår: _____

Kjønn Kvinne Mann

Bostedskommune: _____

Deltakerkode:

Del B: Språklig bakgrunn

Morsmål

Er norsk morsmålet ditt?

Ja Nei

Hvis ja, har du andre morsmål i tillegg?

Ja Nei

Hvis ja, hvilke(t) språk? _____

Hvilket språk bruker dere hjemme? _____

Hvor ofte leser du tekst skrevet på norsk?

Hver dag Flere ganger per uke Et par ganger i uken Av og til Aldri

Hvor ofte skriver du tekst på norsk?

Hver dag Flere ganger per uke Et par ganger i uken Av og til Aldri

Engelsk og andre fremmedspråk

I engelsk, hvordan vurderer du ferdighetene dine på hvert av disse områdene?

	Grunnleggende	Middels	Avansert	Flytende
Lesing				
Skriving				
Snakke				
Lytte				
Totalt				

Har du bodd i, eller hatt lengre opphold i, et land hvor engelsk er hovedspråk?

Ja Nei

Hvis ja, hvor lenge varte oppholdet/oppholdene? _____

Har du vært på kortere (under 14 dager) reise i et land hvor engelsk er hovedspråk?

Ja Nei

Har du bodd i, eller hatt lengre opphold i, et land hvor annet enn engelsk er hovedspråk?

Ja Nei

Hvis ja, hvor var det, og hvor lenge varte oppholdet/oppholdene? _____

Hvilke språk kan du utover morsmålet ditt og engelsk?

Språk	Nivå			
	Grunnleggende	Middels	Avansert	Flytende
Tysk				
Fransk				
Spansk				
- angi språk				
- angi språk				
- angi språk				

Hvor ofte leser du tekster på engelsk?

- Hver dag Flere ganger pr uke Et par ganger i uken Av og til Aldri

Hvor ofte skriver du tekster på engelsk?

- Hver dag Flere ganger pr uke Et par ganger i uken Av og til Aldri

Hvor ofte lytter du til/hører du engelsk?

- Hver dag Flere ganger pr uke Et par ganger i uken Av og til Aldri

Hvor ofte ser du engelskspråklige serier/filmer?

- Hver dag Flere ganger pr uke Et par ganger i uken Av og til Aldri

Når du ser engelskspråklige filmer, hvilken av disse alternativene bruker du oftest?

- Undertekst på norsk (morsmål) Undertekst på engelsk Ingen undertekst

Hvor ofte ser du engelskspråklige tegneseriefilmer/serier?

- Hver dag Flere ganger pr uke Et par ganger i uken Av og til Aldri

Har du sett tegneserien “Family Guy”?

- Ja Nei

Hvis ja, i hvor stor grad? _____

Hvor ofte spiller du engelskspråklige dataspill?

- Hver dag Flere ganger pr uke Et par ganger i uken Av og til Aldri

Hvilken type spill spiller du? _____

Hvor mange timer per dag? _____

Hvor mye TV ser du hver dag?

- 7 timer eller mer 5-6 timer 3-4 timer 1-2 timer Ser ikke TV

Del C: Andre faktorer i språklæring

Har du, eller har du hatt, problemer med synet utover normal brillebruk?

Ja Nei

Har du, eller har du hatt, problemer med hørselen?

Ja Nei

Har du, eller har du hatt, språkvansker av noe slag (spesifikke språkvansker, lese-/lærevansker eller lignende)?

Ja Nei

Hvis ja, spesifiser: _____

Har du, eller har du hatt, andre diagnoser som kan tenkes å påvirke språklæring (ADHD, autisme eller lignende)?

Ja Nei

Er du venstrehendt?

Ja Nei

Del D: Vokabulartest og grammatikktest

Resultat vokabulartest:

Runde 1.

Runde 2.

Resultat grammatikktest:

Appendix C – Comprehension questionnaire

Note: X= correct response

Deltakerkode: (Fås av læreren)
--

Select the correct alternative. Select only one alternative for each question.

1. Jillian's boyfriend is...

- a. Stewie
- b. Peter
- c. Brian X
- d. Carl

2. Stewie is...

- a. A dog
- b. A girl
- c. A boy X
- d. A cat

3. Brian is...

- a. A dog X
- b. A girl
- c. A boy
- d. A cat

4. "Pe-ople" is...

- a. A television show
- b. A magazine X
- c. A movie
- d. A girl

5. Peter and Jillian go to see...

- a. Les Miserables
- b. Walt On Ice
- c. Disney On Snow
- d. Disney On Ice X

6. The book Brian is writing is called...

- a. Faster than the speed of love X
- b. Faster than the speed of lightning
- c. Fast and furious
- d. Faster than the speed of ice skaters

7. Lois...

- a. Is in love with Brian
- b. Likes Brian's book idea
- c. Wants to direct a movie based on Brian's book
- d. Makes fun of Brian's book idea X

8. Stewie is looking for...

- a. French fries
- b. Fanta lemon
- c. Graham crackers X
- d. Chocolate cake

9. Jillian's neighbour is...

- a. A sailor
- b. A pirate X
- c. An opera singer
- d. A thief

10. Jillian ends the relationship because...

- a. Her boyfriend is too lazy
- b. Her boyfriend lied about wanting to live with her X
- c. Her boyfriend wants to move to Africa
- d. Her boyfriend kissed another girl

11. Stewie has...

- a. An Australian accent
- b. An Indian accent
- c. An American accent
- d. A British accent X

12. Meg goes to buy...

- a. Diapers X
- b. Deodorant
- c. Underwear
- d. Magazines

13. Carl is...

- a. A hair-dresser
- b. A shop manager X
- c. A sports commentator
- d. A movie star

14. Chris is Meg's...

- a. Baby
- b. Boyfriend
- c. Bingo partner
- d. Brother X

15. Chris tries to pay for comic books with his...

- a. Money
- b. Insects
- c. Pooh X
- d. Pee

16. Carl and Chris discuss...

- a. TV shows
- b. Christmas
- c. Cover girls
- d. Movies X

17. Meg eventually gets a job for...

- a. A television show
- b. A phone-sex line X
- c. A telemarketing company
- d. A travelling circus

18. Chris helps Meg by...

- a. Getting her job back X
- b. Saving her life
- c. Buying her jewelry
- d. Lending her money

Appendix D – Word definition task

Note: X = correct response

Deltakerkode:

Select the most appropriate definition for each word. Select only one alternative for each word.

1. A novel

- a. A short piece of narrative prose fiction
- b. A fictional prose narrative of considerable length X
- c. A prose narrative consisting of minimum two volumes
- d. An encyclopedia

2. Rent

- a. Money used to pay back a bank loan
- b. Payment for living in an apartment/house which belongs to someone else X
- c. An object free from dirt, stain, or impurities
- d. Fixed compensation for services, paid to a person on a regular basis

3. To hire

- a. To get rid of an employee
- b. To tie a metallic rod
- c. To employ someone X
- d. To advance to a higher position

4. A supervisor

- a. Someone who works in a supermarket
- b. An employee
- c. A person responsible for cleaning the bedroom
- d. Someone in charge of a particular unit X

5. To promote

- a. To raise someone to a higher rank X
- b. To assign someone to a lower position
- c. To give someone a second chance
- d. To prefer someone to another

6. A sequel

- a. A constitution
- b. A continuation X
- c. Someone who sneaks about
- d. A conclusion

7. To smother

- a. To cough
- b. To bore
- c. To die
- d. To suffocate X

8. A pamphlet

- a. A short piece of printed paper on a current topic X
- b. A porcupine
- c. A soft piece of toilet paper
- d. A flying piece of printed paper on a current topic

9. Off-putting

- a. Turning something off
- b. Interesting
- c. Cheating
- d. Repelling X

10. Hazy

- a. Hazardous
- b. Undecided
- c. Unclear X
- d. Unsited

11. Disposable income

- a. The money someone pays in taxes
- b. Diapers that can only be used once
- c. The money someone has available to buy consumer goods X
- d. The money someone has to borrow to afford to buy consumer goods

12. A cruller

- a. Sweet cake dough fried in deep fat X
- b. French fries
- c. A fruit-flavored and glazed twisted candy roll
- d. A hair curler

13. Crabby

- a. Humorous
- b. Ill-tempered X
- c. Lazy
- d. Annoying

14. The gist of something

- a. A central idea X
- b. A protagonist
- c. A conception
- d. A poison

15. Braces

- a. A dental regulating tooth brush
- b. An appliance that creates dental irregularities
- c. A dental device that is used for flossing
- d. An appliance that corrects dental irregularities X

16. An underrated actress

- a. A female with low self-esteem
- b. A female whose acting talent is underestimated X
- c. An underestimated female magician's assistant
- d. A female with great acting talent who has low self-esteem

17. A captive

- a. A person who only writes in capital letters
- b. A person who has committed a crime
- c. A person who is held against his will X
- d. A person who is held responsible for an accident

18. To stack

- a. To arrange things in a pile X
- b. To organize things in a line
- c. To hit someone with a stick
- d. To arrange piles of hay in a line

19. “No way”

- a. “You are joking”
- b. “Not possibly” X
- c. “Not a road”
- d. “Never mind”

20. To alphabetize

- a. To learn the alphabet
- b. To arrange in alphabetical order X
- c. To mark with letters
- d. To organize by date

21. Paper route

- a. The job of delivering newspapers regularly X
- b. The road taken by the toilet paper after flushing
- c. The job of delivering mail regularly
- d. The process in which a newspaper is made

22. Committed

- a. To be happy about a change
- b. To leave someone out
- c. To be bound to something X
- d. To evolve into something more serious

Idioms

23. To be game

- a. To be happy
- b. To be agreeable to participate in something X
- c. To be excited about an upcoming event
- d. To be willing to play a game

24. To take charge

- a. To be responsible
- b. To reload batteries
- c. To take control over something X
- d. To make someone pay for something

25. To make up one's mind

- a. To conclude a chapter in a book
- b. To choose the direction of a play
- c. To paint one's face
- d. To come to a decision X

26. A pain in the ass

- a. A suppository pill
- b. Something causing trouble X
- c. A donkey with problems
- d. Constipation

27. To cheer up

- a. To buy new furniture
- b. To view something from above
- c. To become happy X
- d. To sing cheerfully

28. To get pushed around

- a. To do something bad
- b. To be supported by someone
- c. To be physically harassed
- d. To do everything you are being told X

29. To have a broken heart

- a. To experience great emotional pain X
- b. To have a cardiac arrest
- c. To feel a strong pain in the chest
- d. To be broke

30. Not my cup of tea

- a. A borrowed cup used for drinking tea
- b. A type of tea you do not like
- c. Something you enjoy or like
- d. Something you do not like X

Appendix E – Lexical decision task

Note: o-occurred, n-did not occur, p-primed word, v-visually primed word

<input type="checkbox"/> Pilot	o	<input type="checkbox"/> Begin	p
<input type="checkbox"/> Hot	o	<input type="checkbox"/> Aerial	o
<input type="checkbox"/> Pirate	v	<input type="checkbox"/> Immigrant	o
<input type="checkbox"/> Pressed	p	<input type="checkbox"/> Stunning	n
<input type="checkbox"/> Flat	n	<input type="checkbox"/> Money	o
<input type="checkbox"/> Equal	o	<input type="checkbox"/> Accuse	n
<input type="checkbox"/> Overpaid	o	<input type="checkbox"/> Stock room	o
<input type="checkbox"/> Error	n	<input type="checkbox"/> This is functioning	p
<input type="checkbox"/> Job	o	<input type="checkbox"/> Biscuit	p
<input type="checkbox"/> Dumb	n	<input type="checkbox"/> Ugly	n
<input type="checkbox"/> Ache	n	<input type="checkbox"/> Be sorry	o
<input type="checkbox"/> Will do	p	<input type="checkbox"/> Pig	p
<input type="checkbox"/> Work	o	<input type="checkbox"/> Paycheck	o
<input type="checkbox"/> Bedroom	v	<input type="checkbox"/> Cupboard	n
<input type="checkbox"/> Backpack	p	<input type="checkbox"/> Sexy	p
<input type="checkbox"/> Daughter	n	<input type="checkbox"/> Skating duet	o
<input type="checkbox"/> Fatherly	o	<input type="checkbox"/> Dog	v
<input type="checkbox"/> Kill	n	<input type="checkbox"/> Constitution	n
<input type="checkbox"/> Genuine	p	<input type="checkbox"/> Awesome	o
<input type="checkbox"/> Move in	o	<input type="checkbox"/> State	n
<input type="checkbox"/> Basketball	n	<input type="checkbox"/> Brand-new	o
<input type="checkbox"/> Film	o	<input type="checkbox"/> Motherly	n
<input type="checkbox"/> Married	n	<input type="checkbox"/> Lonesome	n
<input type="checkbox"/> Sabbatical	o	<input type="checkbox"/> Paper towel	o
<input type="checkbox"/> Shop	v	<input type="checkbox"/> Chick	o
<input type="checkbox"/> Cry	n	<input type="checkbox"/> Dreadful	p
<input type="checkbox"/> Cobbler	o		