Dissertation in part-fulfilment of the MSc in Applied Linguistics and Second Language Acquisition
August, 2022

The effect of English videos with different types of subtitles and glosses on vocabulary learning of Chinese ESL learners

Word Count: 17110

Yaguang Li

Abstract

There is growing interest in research focusing on the effect of subtitled videos on second language (L2) vocabulary learning. While a large number of studies have investigated the benefits of L2 subtitles, few studies have examined L1 subtitles and dual subtitles and compared the effects of the subtitles in different languages. What's more, glosses in subtitles are paid more attention to, because they stimulate learners' attention to the novel words and provide the meaning of the words to scaffold acquisition of form-meaning links. The present study investigated the effects of five different types of subtitles (L2 subtitles, L1 subtitles, dual subtitles, L2 subtitles + L2 glosses, and L2 subtitles + L1 glosses). A total of 314 Chinese middle school students who are beginner learners of English as a second language (ESL) were randomly assigned to one of the five conditions and watched two videos with different types of subtitles. For each video, 10 target words were included, and there were 20 target words in total. An unexpected post-test measuring the participants' word form and word meaning learning was administered immediately after the participants watched each video. Results show that the group with L2 subtitles and L1 glosses significantly outperformed the L1 subtitles group in the word form tests, while there was no significant difference between any other groups. For word meaning learning, it was found that both the L1 subtitles group and the dual subtitles group significantly outperformed the L2 subtitles group. What's more, the group with L2 subtitles and L1 glosses significantly outperformed every other group in the word meaning test. However, the group with L2 subtitles and L2 glosses had no significant difference from the L2 subtitles group without glosses. This interventional study contributes to the literature by providing evidence which suggests that L1 subtitles may benefit word meaning learning, and that glosses can draw learners' attention to the target words and scaffolds word form learning. Pedagogical implications on the use of subtitles and glosses in L2 learning videos are provided.

Acknowledgements

First and foremost, I want to express my greatest gratitude to my supervisor, Dr. Sophie Booton, who have always been responsive and supportive throughout the whole year. She has provided helpful guidance not only to the research design and the data analysis, but also to academic writing. I have learned so much from her and this study can never be so great without her help.

I want to thank my parents who have been always supporting every decision I made. I want to thank my mother as well as her friends, who are all great teachers and helped me recruit the participants. I could never recruit such a large number of participants without their help.

I also want to express my gratitude to everyone in the MSc ALSLA group. I want to thank the lecturers who delivered valuable knowledge to us. I also want to thank all my friends who provided me with great help throughout the year. Besides, I also want to thank the members from the REAL group, who have provided valuable and helpful suggestions on my research design.

At last, I want to thank the teachers who helped distribute the information and the participants who spent time to complete this intervention online. Their contribution to this study is very valuable.

List of Figures

List of Tables

Contents

Chapter 1 Introduction

Vocabulary is an essential component of language learning, especially for students learning English as a second language (ESL). However, there are a large number of words that need to be learnt to comprehend English texts or audios: 6000-7000 word families are required for listening, and 8000-9000 are required for reading (Nation, 2006). In this case, learning L2 vocabulary intentionally in formal classroom instructions may be insufficient for learners to acquire so many words. Therefore, more opportunities to learn the words incidentally outside the classroom are necessary for ESL learners, especially for learners who are living in a country that does not speak the language and are not immersed in authentic language contexts. That is to say, incidental vocabulary learning could be an important method for the learners to acquire more new words out of the classroom.

Incidental vocabulary learning is one of the major ways to acquire vocabulary in a second language (L2). Incidental vocabulary learning involves picking up the vocabulary through language related activities (Barcroft, 2009; Hulstjin, 2003), which is opposed to intentional vocabulary learning which refers to deliberately memorizing the new words (Nation, 2001). By contrast, incidental vocabulary learning does not entail any deliberate memorization of the target words. Incidental and intentional are two tendencies to learn L2 vocabulary: while intentional vocabulary refers to learning new words with the intention to do so, incidental vocabulary learning refers to the opposite. Most vocabulary learning activities could be placed at a point on the continuum between purely intentional and purely incidental, the two theoretical extremes that are unlikely to see in real life (Coady & Huckin, 1997; Gass, 1999). Incidental vocabulary learning activities often require language related inputs, including written texts, audios, and videos. Among these different types of input, L2 videos are popular in incidental vocabulary learning as language learners are having more and more access to online materials, including online videos (Vanderplank,

2010). The audio-visual input provided by videos could facilitate L2 language processing through providing visual cues (Brett, 1995), and are found to be more beneficial to vocabulary acquisition when compared with audio materials (Brett, 1998; Meskill, 1996).

A problem with the L2 videos, though, is that they may often be too difficult for learners, especially for those in low to intermediate proficiency levels who have not gained enough word families for comprehending the audio-visual input (Danan, 2004; Nation, 2006). It may be difficult for them to identify the forms of the novel words and infer their meanings from the context. Subtitles, which are synchronous on-screen texts accompanying the videos, could possibly enhance video comprehension and vocabulary learning and have received wide attention in the academic world (Montero Perez, 2022; Montero Perez et. al., 2013; Teng, 2020b; Wei & Fan, 2022).

There are many different types of subtitles that have been researched. First, there are subtitles in different languages. L2 subtitles, often called captions in previous studies, are the L2 transcripts of the videos that are synchronously shown on the screen. Although the terms “caption” or “L2 caption” are frequently used in studies that only focused on L2 subtitles, this study will use the term “L2 subtitles” instead because there are many different conditions in this study and using the word “caption” may cause confusions. L1 subtitles are the L1 translation of the contents of the videos that are presented synchronously with the audio. Dual subtitles refer to the combination of L1 subtitles and L2 subtitles, i.e., showing both L2 transcripts and the corresponding L1 translation at the same time (Dizon & Thanyawatpokin, 2021). However, these subtitles in different languages do not stimulate learners’ noticing of novel words in the input, and therefore learners may not always form their form-meaning links of the words (Laufer, 2003, 2005).

Hence, in order to stimulate learners’ noticing toward target words and scaffold form-

meaning link acquisitions, different ways to provide visual salience in subtitles have been examined, mostly used in L2 subtitles (Abraham, 2008). For example, keyword subtitles do not provide the full transcript, but only the keywords that may be difficult for the learners to understand. Full subtitles with highlighted keywords still provide the full subtitles but highlight the difficult keywords instead. However, although these two types of subtitles can stimulate learners' attention towards the novel words, the learners may not acquire all the form-meaning links as the word meaning is not directly provided and sometimes learners may not be able to infer the word meaning from contexts.

Glosses are another type of salience in subtitles that provide the meaning of unfamiliar words. Glosses are the interpretation or translation of unfamiliar words that are located in the side or bottom margins of the input (Lomicka, 1998). They can not only highlight the difficult words to make the learning process more intentional, but at the same time provide meanings of the words and therefore scaffold learning of form-meaning links as well (Montero Perez et. al., 2018).

Figure 1 shows what different types of subtitles may look like when L1 is Chinese and L2 is English. The sentence is "The horse is galloping across the land", and "gallop" is the keyword which may be novel to the learners. The figure includes L2 subtitles, L1 subtitles, dual subtitles, L2 keyword subtitles, full L2 subtitles with highlighted keywords, and L2 subtitles with L1 glosses. Although studies have compared some types of subtitles (Montero Perez et. al., 2014, 2018; Teng, 2018, 2019c, 2019a, 2020), few studies have examined the differences of subtitles and glosses in different languages.

Figure 1 Different Types of Subtitles.

This study focuses on L2 vocabulary learning through viewing subtitled videos and investigates the effects of five different types of subtitles featuring different languages and glosses. First, three types of subtitles in different languages are included, i.e., L2 subtitles, L1 subtitles, and dual subtitles. Second, two types of subtitles with glosses in different languages are included, both of which are based on L2 subtitles: they are L2 subtitles + L1 glosses and L2 subtitles + L2 glosses. This experimental study aimed to provide insight into the effect of using different types of subtitled videos for middle school ESL learners.

Chapter 2 Literature Review

2.1 Introduction

This chapter will review the theoretical background and empirical studies on subtitled videos and their influence on L2 incidental vocabulary learning. There are different types of subtitles used in foreign language videos to boost L2 learning. L2 subtitle, also caption, the literal transcript of the contents in the video, is one of them. However, there are also many other types of subtitles, including those in different languages and those used together with other literal cues like glosses. This chapter will first provide the theoretical background of using subtitled videos and will introduce empirical studies that focus on use of subtitles in L2 vocabulary learning. The use of subtitles in L2 vocabulary learning in general will first be reviewed, and then studies that focus on different types of subtitles will be looked at, including subtitles in different languages and subtitles with glosses. The identified research gaps and the research questions for the current study will also be presented at the end of the chapter.

2.2 Theoretical background for using subtitled videos in L2 language learning

To begin with, theories have argued that videos which combine visual and audio information are more beneficial than audio input. With the fast development of technology, combining audio materials with visual supports in videos has been widely adopted as a means of L2 learning (Baltova, 1994; Vanderplank, 2010). The use of video showed better effects in language learning when compared with traditional audio input, because the visual cues provided by the video could scaffold understanding (Montero Perez et. al., 2013). This is especially true for the learners who are not proficient enough and cannot process the audio cues automatically, and

the visual aid can help them better understand the contents. Meskill (1996) argued that the visual material could “enrich target language processing” (p.196). Other studies have also investigated the use of multimedia in incidental language learning and have implied that the use of video is able to improve vocabulary acquisition when compared with audio materials (Brett, 1998; Duquette and Painchaud, 1996).

After suggesting the advantage of using videos for language learning, more studies focused on how to modify the videos to make the vocabulary learning process more efficient. The use of subtitles, i.e., a stream of written text presented synchronously with the video (Danan, 2004), emerged as one of the major ways to modify the videos. It is argued that the combination of videos and texts could provide the learners with a bi-modal input, which can provide the learners with multiple representations of the same information (Vanderplank, 2016). It has been widely argued that the use of subtitles in L2 learning videos could facilitate comprehension (Montero Perez et. al., 2013; Teng, 2019a) and vocabulary learning (Montero Perez et. al., 2013, 2018; Teng, 2019c, 2019b, 2020, 2022).

There are three major theoretical frameworks that explain and support the use of subtitles in L2 incidental vocabulary learning, all of which focus on information processing: Paivio’s (1986) dual-coding theory, Baddeley’s (1986) working memory theory, and Mayer’s (2001) multimedia principle. They will be discussed in detail in the following section.

2.2.1 Dual-coding Theory

Dual-coding theory is based on the view that cognition consists of symbolic representational systems that incorporate both verbal and nonverbal knowledge. That is to say, human beings can process language and nonverbal cues simultaneously. Based on this general view, dual-coding theory hypothesized that there are two separate subsystems for the representation and processing of nonverbal cues and

language respectively, which are both structurally and functionally distinct (Sadoski et. al., 1991). This hypothesis assumes that the language system and the non-verbal system differ in the way they are structured, and they are independent from each other, which means they can be active at the same time, or one can be active without the other (Paivio, 1986). On the other hand, they are interconnected functionally, indicating that activating one of them may trigger the activity of the other.

When applied to subtitled videos, it is argued that such videos are able to present the same thing in both verbal and nonverbal way. The subtitles, together with the speech, could activate the learners' verbal system, while the graphics could trigger the nonverbal imagery system. When the two systems are activated at the same time, information processing could be improved with deeper processing and better recall, enabling the learner to process the verbal-visual connection, and to store the information (Montero Perez et. al., 2013; Park, 2004). Subtitles may help the learners to retrieve a word and the visual cue at the same time, and therefore scaffold learning the form-meaning link of a new word.

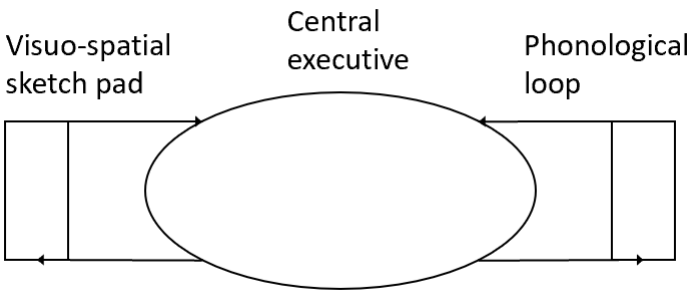
2.2.2 Working Memory Model

As shown in Figure 2, Baddeley's (1986) working memory model consists of three core components: central executive, and two subordinate memory systems: visuo-spatial sketch pad and phonological loop. The central executive system is able to control the memory, and it coordinates information retrieved from the two subsystems which are responsible for visual and phonological coding respectively. The visuo-spatial sketch pad is a part of visual coding, and it handles spatial and imagery information. The phonological loop, on the other hand, is a component of acoustic and phonological coding, including deciphering speech-based information, including vocabulary (Baddeley, 1992).

Similar to the dual-coding theory, the working memory model assumes two separate

systems, responsible for visual and aural cues respectively. The difference is that the working memory model have a central executive system which coordinates and controls the storage subsystems, and in this model, the subsystems store visual and aural information respectively instead of verbal and non-verbal information. When the learning is watching a subtitled video, the visual and aural information about the same word can be presented simultaneously, and the learners' working memory about both visual and aural information can be processed by the central executive, and therefore long-term acquisition of form-meaning links can be enhanced. Both theories suggest that subtitles could enhance the knowledge processing of learning a new word through activating both subsystems at the same time.

Figure 2 Baddeley's (1986) working memory model.



2.2.3 Multimedia Theory

Mayer's (2001, 2005) theory on multimedia learning is based on Paivio's (1986) dual-coding theory and targets at language learning. This theory was raised in the context of learning the first language and is about how multimedia with multiple input sources may influence language learning. The multimedia principle is a part of the theory, which suggests that "people learn better from words and pictures than from words alone" (Fletcher & Tobias, 2005, p. 117). Based on the dual-coding theory that the visual and verbal processing systems are two separate systems, this principle implies that integration of texts and graphics of the words could result in

faster and deeper processing through stimulating the visual and verbal systems at the same time.

However, not every aspect of Mayer's cognitive theory of multimedia learning is in favour of using the simultaneous representation of texts, sounds and images. In Sweller's (2005) redundancy principle, it was argued that instead of facilitating learning, the simultaneous presentation of the same information in different formats (audio, imagery, textual) could interfere with learning new words. It was argued that redundant representations of the same information could exert great cognitive load on the learners and lead to unnecessary burdens of working memory (Bobis et. al., 1993). Working memory resources may be allocated to coordinate information from difference sources instead of being used for learning and processing the inputs (Sweller, 2005). Besides, as these theories were based on first language learning, it is still not sure whether they are applicable to second language learning.

2.3 Empirical studies on L2 Incidental Vocabulary Learning from Subtitles

While empirical studies have been done to examine different types of subtitles, most studies investigated the effect of L2 subtitles, i.e., the captions that synchronously show the contents of the audio, and a large number of studies have suggested the effectiveness of L2 subtitled videos in vocabulary learning compared to non-subtitled videos (Bird & Williams, 2002; Chai & Erlam, 2008; Markham, 1999; Montero Perez et. al., 2015, 2018; Sydorenko, 2010; Teng, 2019a, 2022; Winke et. al., 2010, 2013). Montero Perez et. al. (2013) did a systematic review with a meta-analysis of all previous studies addressing the effectiveness of L2 subtitled video on incidental L2 vocabulary learning. They identified 18 studies, 10 of which were included in the meta-analysis. The 10 studies included participants from different backgrounds who speak different languages: Asian, English and Spanish as L1 speakers were recruited,

and they spoke English, Spanish, French, and Russian as a second language. Besides, participants with different proficiency levels: 6 studies included beginner learners, 6 studies included intermediate learners, and 2 studies included advanced learners. However, the diversity in age is relatively small: 7 studies recruited adult learners from universities, and there were no studies that focused on the younger learners. According to these 10 studies, it was calculated that the mean effect size of L2 subtitling on an immediate vocabulary post-test was large ($g = 0.87$), and students who watched videos with L2 subtitles significantly outperformed students in the control groups ($p < 0.001$), both in recognition tests and recall tests. Besides, it was found that learners in all proficiency levels benefited from L2 subtitling. This meta-analysis indicates that L2 subtitled videos are better than non-subtitled videos in supporting L2 vocabulary learning in the short term, including both form and meaning learning. However, there have been fewer studies that focus on young learners.

Sydorenko's (2010) study is one of the studies that examined the effect of the multi-modal input of L2 subtitled videos on L2 vocabulary learning. This study recruited 26 Russian as second language learners in a university and assigned them into three conditions with three different types of videos: video with audio and L2 subtitles (captions) (VAC), video with audio (VA), and video with L2 subtitles (captions) (VC). The participants completed word form recognition and word meaning translation test in both written and aural forms. The results indicate that both groups with L2 subtitles scored higher on the written test than the aural test, and that the VA group scored higher on the aural test than the written test. This suggests that although L2 subtitles provide support for the written forms of the words, they might detract the learners' attention from the aural aspects of language. Besides, learners in the VAC group outperformed the VA group in the word meaning test. It was argued that L2 subtitles facilitate recognition of written word forms, and the combination of L2 subtitles and audio in a video could scaffold better word meaning learning. What's

more, according to an additional questionnaire, participants paid more attention to the L2 subtitles than the video or audio. This study adds evidence to support the effectiveness ofL2 subtitles in videos and is innovative in that it uses both written and aural forms of tests. However, this study had a small sample size, which weakens the results. Besides, although the study investigated their viewing strategies of L2 subtitled videos, the strategies were self-reported by the learners, which may not by reliable. Therefore, studies that include statistical analyses on the learners' strategies of allocating attention were still needed.

In order to further explore how learners allocate their attention when watching subtitled videos, eye-tracking studies were done to try to probe into how L2 subtitles work in learning vocabulary and how learners process the simultaneous intake of audio, video and test (Winke et. al., 2013; Montero Perez et. al., 2015). For example, Winke et. al.'s study (2013) investigated how often learners read the L2 subtitles, and whether the specific foreign language being learnt, and the content familiarity may influence the effect of L2 subtitled video. This study recruited 33 college students who have English as their first language and are learning Arabic, Chinese, Russian, and Spanish as a foreign language. They were presented two different videos, one concerning a familiar topic and another containing novel content. The videos were originally in English, and were transcribed, translated and dubbed into the four target foreign languages. The participants' eye movements were tracked, and they finished a comprehension test for each of the videos. The results show that on average, the participants read the L2 subtitles 68% of the time any subtitles are presented on the screen. This indicates that L2 subtitles are attracting major attention, which is consistent with previous results that learners pay more attention to the L2 subtitles than the video or audio (Sydorenko, 2010). Besides, the number is a lot less than 100%, which indicates that the learners are also paying attention to the imagery contents of the video, and watching subtitled videos is different from reading the texts directly. The results indicate that while learners pay much attention to L2 subtitles

and L2 subtitles do facilitate L2 vocabulary learning, the effect of L2 subtitles may vary across different languages and materials.

From the review above, it can be found that abundant studies have validated the facilitating effect of L2 subtitled videos in L2 vocabulary learning compared with videos without subtitles. It was found that L2 subtitles could enhance both word form and word meaning learning. What's more, L2 subtitles are effective for participants in different proficiency levels and of different ages. Apart from that, eye-tracking studies indicate that learners pay major but not all attention to the L2 subtitles in the video. The findings of the empirical studies are consistent with the predictions according to the information processing theories that predicted the effectiveness of subtitles in L2 learning videos.

However, according to a recent literature review on the studies on on-screen texts in audio-visual input for L2 vocabulary learning, it was pointed out that in the past decade, more studies were comparing different types of subtitles, and have tried to find the most beneficial type of on-screen text (Wei & Fan, 2022). That is to say, few recent studies looked at only the effect of L2 subtitles. Different types of on-screen texts were categorized and compared. Therefore, the following reviews will discuss the types of subtitles that this study focuses on, and previous studies that compared their effects.

2.4 Subtitles in different languages

2.4.1 L1 subtitles

While many studies have proved the effectiveness of L2 subtitles, i.e., captions on incidental vocabulary learning in the second language, fewer studies have examined whether L1 subtitles may have a similar effect when learning a second language. L1

subtitles provide the viewers with the translation of the audio input in the foreign language into their first language (Almeida & Costa, 2014), and could be useful because it provides the meaning of the contents and may facilitate acquisition of form-meaning links.

Empirical studies on the effect of L1 subtitles were mainly conducted in the 1990s and on TV programs or movies, rather than videos on the internet (d'Ydewalle & Pavakanun, 1995, 1997; d'Ydewalle & Van de Poel, 1999; Koolstra & Beentjes, 1999). Koolstra & Beentjes' (1999) study is a good example of the studies on L1 subtitled television programs. They recruited 246 Dutch children in Grades 4 and 6 who are beginner learners of English, and investigated their learning of English vocabulary through watching an English video with Dutch subtitles. They were randomly assigned to three conditions: watching an English TV program with L1 subtitles, watching the same English program without subtitles, and watching a Dutch television program as a control group. It was found that the results of written word form recognition and meaning recall tests of the subtitled condition were significantly higher than the other two conditions. The results indicated that L1 subtitles also have a beneficial effect on L2 vocabulary learning, compared to no subtitles. Danan (2004) argued that although learners may automatically read the L1 subtitles, they process the subtitles in parallel with the audio, leading to a similar dual-processing with that when watching a L2 subtitled video. However, these studies did not compare the differences between L1 and L2 subtitles in the interventions.

2.4.2 L1 subtitles versus L2 subtitles

Although both L1 subtitles and L2 subtitles lead to more L2 vocabulary learning than a non-subtitled video, it still needs to be examined whether the effect of L2 subtitles and L1 subtitles are different. It was often predicted that L2 subtitles may be more beneficial for word form learning because it provides the correct written form of the target words, and that L1 subtitles may facilitate word meaning learning more, as it

provides the meaning of the target words. However, according to recent studies that compared these two types of subtitles, the results are mixed and inconclusive. Bisson et. al. (2014) examined whether three different types of subtitles: L1 subtitles, L2 subtitles, and reversed subtitles (L1 sound, L2 subtitles) could facilitate learning Dutch as a foreign language. The results in word form recognition tests did not show any significant difference between the three conditions. The authors explained that this may be due to the word recognition test may not be sensitive enough to measure learning gains, and that the participants did not have any former knowledge in Dutch. This indicates that a certain level in the target language may be required for the audiovisual input to be effective for the learners.

The results of the following studies that compared L1 and L2 subtitles were inconsistent. Vulchanova et. al. (2015) compared the effect of L1 subtitles, L2 subtitles, and no subtitles on Norwegian EFL learners between 16 and 17. The learners were tested knowledge of the word forms and meanings four weeks after watching the video. The results showed that only L1 subtitles affected the results of the word meaning test, which was moderated by age, and L2 subtitles had no significant effect on word meaning learning. Different conditions had no effect on the word form recall test, but is was highly possible that four weeks is too long for the learners to still memorize the words in the video. By contrast, Frumuselu et. al. (2015) did a longitudinal study which had contradictory results from Vulchanova et. al.'s (2015) study. This study examined the effect of L1 Spanish subtitles and L2 English subtitles on EFL university students' vocabulary form and meaning learning. It was found in this study that although both the L2 subtitles group and the L1 subtitles group saw learning gains in vocabulary after watching the videos, the L2 subtitles group had greater gains. Such inconsistency results may be due to the difference in L2 proficiency: learners with lower proficiencies may benefit more from L1 subtitles, while those with higher proficiencies may benefit more from L2 subtitles. Apart from that, different video types and test formats may also be reasons

for the inconsistent results.

Peters et. al. (2016) tried to provide more evidence on the difference of the two types of subtitles with two studies on students in a general secondary school and in a vocational school respectively. The students from the general secondary school were intermediate EFL learners, and those from the vocational school had low proficiency in English. The results show that the L2 subtitles group performed significantly better than the L1 subtitles group in the form recall test, but only for intermediate students with a larger vocabulary size, L2 subtitles group performed better than L1 subtitles group in the form recognition test. The two conditions did not vary in their effects on learning word meaning. Besides, it was found that learners with larger vocabulary sizes learn more words through subtitled videos. It was argued that L2 subtitles can not only draw learners' attention to word forms, but also provide the correct written form, which helps the learners with form recall and recognition (Danan, 2004). However, they may not be so helpful for word meaning. Besides, although the subtitles have beneficial effect on both intermediate learners and low-proficiency learners, vocabulary size have a significant and positive relationship with learning gains. This indicates that the more words a student knows in advance, the more words they may pick up when watching a subtitled video. The following studies showed a similar pattern with the results of this study: some find that L2 subtitles are most beneficial for language learning in both word form and meaning (Peters, 2019), and others did not find significant difference between L2 subtitles and L1 subtitles (Matielo et. al., 2017; Pujadas & Muñoz, 2019). In conclusion, there are controversies in the experimental results on the comparison between L1 subtitles and L2 subtitles.

There are few studies that examined the effect of dual subtitles, i.e., the combination of L1 subtitles and L2 subtitles (Dizon & Thanyawatpokin, 2021; Lwo & Lin, 2012; Raine, 2013; Wang, 2019). What's more, most of them did not answer the question whether dual subtitling is better in facilitating vocabulary learning than L1 subtitles or

L2 subtitles alone: some studies did not find a facilitating effect of any type of subtitles and the videos or tests could be too difficult for the leaners (Lwo & Lin, 2012; Raine, 2013), and Wang's (2019) study is not well controlled as all participants watched a video under each of the conditions, and it is questionable whether the learning outcomes for each different video is comparable. Dizon and Thanyawatpokin's (2021) study is the only existing study that effectively compared the effect of dual subtitles with the two types of subtitles alone. They recruited 96 Japanese university students who are beginner L2 English learners and divided them into three conditions. The participants watched English videos with L1 subtitles, L2 subtitles, and dual subtitles respectively, and finished pre- and post-tests in word form recognition and meaning recall of the target vocabulary. It was found that L1 subtitles and dual subtitles groups outperformed the L2 subtitles group in the tests, indicating that L1 subtitles, either combined with L2 subtitles or alone, are more effective than L2 subtitles alone. The result of this study contradicts with previous studies that foundL2 subtitles more effective than L1 subtitles (Peters, 2019; Peters et. al., 2016). This may be because the learners of this study are beginners, who may not be able to understand the contents of the videos without difficulty and to deduct the meaning of the new words without the support of L1 subtitles.

The recent eye-tracking study conducted by Wang & Pellicer-Sánchez (2022) provided more evidence to support the effectiveness of dual subtitles. They recruited 112 intermediate to advanced Chinese ESL learners who watched a documentary in one of four different conditions: dual subtitles, L1 subtitles, L2 subtitles, and no subtitles. They used pre- and post-tests to assess the participants' knowledge of the target vocabulary in the fields of form recognition, meaning recall, and meaning recognition. Results showed that L2 subtitles were more beneficial than L1 subtitles and dual subtitles in word form learning, and dual subtitles were more effective than L2 subtitles for meaning recognition and more effective than L1 subtitles for meaning recall. Besides, the eye-tracking measurements suggested that participants in the dual

subtitles group spent more time reading the L1 translation of the target words. The results showed great potential of dual subtitles in word meaning, and this may be because dual subtitles directly provide form-meaning links of the new words. By contrast, dual subtitles were less effective than L2 subtitles for form recognition, and according to the results of eye-tracking, this may be because the learners focused more on the L1 contents in the dual subtitles and paid less attention to the word form. In conclusion, while dual subtitles may be beneficial for word meaning learning, they do not scaffold word form learning as much.

From the literature reviewed above, it could be found that the comparison between L1 subtitles, L2 subtitles and dual subtitles has inconclusive and mixed results, with only very limited number of studies. Although all three types of subtitles are effective in incidental vocabulary learning, their effects seem to be largely influenced by many other factors, including learners' proficiency and vocabulary size, as well as the type of video and the content of the tests. Therefore, more evidence is needed to provide more empirical evidence on the effects of subtitles in different languages to the existing literature. What's more, most studies were still on learners with at least intermediate proficiency level, and few studies were done on young learners who are beginners in learning the second language.

2.5 Different forms of subtitles

2.5.1 Incidental and intentional vocabulary learning

Although results from previous research provided abundant evidence to support the effectiveness of L1 subtitles and L2 subtitles in L2 vocabulary learning, there is still a potential shortcoming of subtitles that may limit the efficiency of vocabulary learning, that is, subtitles themselves do not contain features that draw learners' attention to target words, and the vocabulary learning process is mainly incidental. It

cannot be assumed that learners will necessarily pay attention to the novel words in the audio-visual input (Laufer, 2003, 2010).

Vocabulary learning activities can be placed at a point on a continuum between incidental and intentional (Barcroft, 2004; Coady & Huckin, 1997; Gass, 1999). Previous studies have suggested that vocabulary learning activities that are more intentional is more efficient than those that are more incidental (Barcroft, 2009; File & Adams, 2010; Laufer, 2003; Qian, 1996; Tajeddin & Daraee, 2013). However, intentional vocabulary learning, like memorizing word lists, cannot provide the use of a word in context. What's more, it is impossible for learners to acquire thousands of words that are required to comprehend texts and audios if every word is learned intentionally. On the contrary, learners could pick up such knowledge when they are incidentally reading books or watching videos in L2 (Webb, 2020). Therefore, it could be more beneficial for the learners if the advantages of incidental and intentional learning can be combined through making incidental vocabulary learning more intentional.

2.5.2 Full L2 subtitles versus salience cues in subtitling

Aiming at making the vocabulary learning process through L2 subtitled videos more intentional, previous studies have examined the effects of salience in the subtitling line. There are two major ways to do so: through keywords or highlighted items in full L2 subtitles. L2 keyword subtitling refers to showing only the target vocabulary in the L2 subtitles. Initially, keywords were expected to deal with the potential textual density problem (Guillory, 1998; Park, 2004). It was argued that learners may pay too much attention to the L2 subtitles and therefore compromise auditory processing, and keywords can reduce the textual density while still providing support for the learners (Garza, 1991). However, it could also draw the learners' attention to the target words. Another way of to provide salience in L2 subtitles is highlighting the keywords in the full L2 subtitles. This has a similar effect of highlighting the target words with

keyword subtitling, but it still provides the full L2 subtitles.

Empirical studies have found that salience in L2 subtitles, no matter whether the full subtitles are provided, is effective for advanced learners. Montero Perez et. al. (2014) recruited 133 undergraduate students who have Dutch as their L1 and are high-intermediate learners of French as the second language. They were assigned to one of four conditions and watched three French clips twice with different types of L2 subtitles: no subtitle, full L2 subtitles, L2 keyword subtitles, and full L2 subtitles with highlighted keywords. They completed form recognition, meaning recognition, meaning recall, and clip association tests. The results revealed that the three subtitling groups significantly outperformed the control group with no subtitles in form recognition and clip association tests, and there is no significant difference between the three groups with subtitles. However, only two groups with salience in subtitling, i.e., L2 keyword subtitles and full L2 subtitles with highlighted keywords groups achieved significantly higher scores in meaning recognition than the control group. The conditions had no influence on meaning recall. The results suggest that visual salience of target words may draw the learners' attention to the target words and facilitate the process of inferring word meaning.

In another similar study that compared the effect of full and L2 keyword subtitles on high-intermediate learners of French as L2, it was found that learners in the L2 keyword subtitles groups outperformed the full L2 subtitles groups in the form recognition test, but no significant difference was found in clip association and meaning tests (Montero Perez et. al., 2015). The results of these two studies suggest that although visual salience has the potential to draw the learners' attention to the word forms and elaborate their processing of target words, the inconsistent results on the effects of measures on word form recognition tests and word meaning recognition tests suggest such benefits may be limited. The authors explained that the inconsistency may be due to the number of viewings. Learners in Montero Perez et.

al. (2014) watched the clips twice, while those in the 2015 study only watched once. It is possible that when watching the second time, the learners could allocate more attention to the target words in the salience groups, and they can infer the meaning of the words from the context. On the contrary, when the learners watched only once, they may pay more attention to the comprehension of the video and are not available for the inferring process. To sum up, L2 keyword subtitles and highlighting keywords in full L2 subtitles do draw the learners' attention to target words, but they do not directly provide form-meaning links and learners have to infer word meaning from contexts.

While keyword subtitles showed greater benefits than full subtitles for advanced learners with high to intermediate proficiency levels, it is not the same case with young learners with lower proficiency levels. Teng (2019c) focused on measuring the effects of L2 subtitle types (full L2 subtitles, L2 keyword subtitles, and no subtitle) and frequency of target word exposure on L2 vocabulary learning. Participants were 257 Grade 6 primary school students who spoke Chinese as L1 and English as L2. Different from previous studies, the participants are young leaners who are beginners in learning the second language and had been learning the language for five to six years. The results revealed that the group which viewed videos with full L2 subtitles significantly outperformed the L2 keyword subtitles group and the no subtitle group in both word form and meaning tests. Similarly, in a later study, Teng (2019a) examined the effects of the three types of subtitling and an advance-organizer strategy before watching the videos on incidental learning of collocations. The participants were 361 Chinese primary school students with English as their L2, and they watched four videos and were tested with their productive and receptive knowledge of form and meaning of the target collocations. The results showed that full L2 subtitle benefited learning of collocations more than L2 keyword subtitle and no subtitles, and that administration of the advance-organizer strategy resulted in better learning of collocations. The results of these two studies thus contrast with the former studies on undergraduate students

(Montero Perez et. al., 2014, 2015). While the two studies on undergraduate students suggested keyword subtitles may be more beneficial than full subtitles, the two studies conducted by Teng on primary school students found opposite results. Teng explained that full L2 subtitles provided more information for the learners, from which students could derive the meaning of new words. The information included semantic and syntactic contexts, which may enable learners to process more messages. By contrast, L2 keyword subtitles may not provide enough information for these beginners. Although there weren't any individual studies that directly compare the effect of the two types of L2 subtitles on learners with different proficiency levels, the results of these four studies suggest keyword subtitling could benefit learners with high proficiency levels more, as they can draw their attention directly to the target words and the reduced information could still provide enough support for them. For beginners, however, the keyword subtitles may not provide enough information, while full subtitles could facilitate their comprehension and deeper processing of the contexts. Therefore, both types of L2 subtitles have shortcomings that may interfere with vocabulary learning, and for younger learners who are beginners, full subtitles seem to be more appropriate.

2.5.3 Glossing in L2 subtitles

According to the reviews above, salient subtitles still have potential disadvantages. What's more, even if it does draw learners' attention to the target words, it does not directly provide explicit access to word learning. Word meaning is considered one of the most difficult parts of word knowledge to learn incidentally (van Zeeland & Schmitt, 2013). As explained in Teng's (2019a) study, learners may have to infer the meaning of unknown words from the contextual clues in semantics and syntax. However, it could be extremely challenging for the learners, especially for the beginners who have difficulty understanding the video and have to allocate much attention to the contents. Unlike textual input, the real-time nature of video input does not enable the learners to spend much time to derive meaning from contextual

information. Therefore, providing explicit access to word meaning may help the learners make the initial form-meaning links while watching the videos.

Glossing is one of the major ways to provide direct access to meaning when the learners are watching videos. It refers to the interpretation or translation of unfamiliar words that are located in the side or bottom margins of the input (Lomicka, 1998). It not only provides the correct meaning of the new word, but also does not interrupt the input process or require extra time for learning (Ko, 2005).

The use of glosses in incidental vocabulary learning has been investigated in the context of written input, and many studies have lent support for its benefits (Abraham, 2008; Ertürk, 2016; Hong, 2010; Hulstjin et. al., 1996; Jacobs et. al., 1994; Ko, 2005, 2012; Nagata, 1999; Yoshii, 2006). For example, Hulstijn et. al. (1996) recruited 78 Dutch university students who learn French as their L2. The participants read short stories with the help of glosses (L1 translations of unknown words), dictionary (access to a bilingual dictionary), or no help. It was found that both glosses and dictionary enhance learning of form-meaning recognition, but glosses have even better effects than dictionary. It was argued that readers may ignore unknown words, or directly infer the meaning of the words wrongly when they are provided with dictionaries. By contrast, glosses can draw learners' attention to unknown words enhance form-meaning links when reading the texts. This study showed the great potential of glosses in the written texts.

Similar to subtitles, glosses in different languages may have different effects on L2 vocabulary learning. L1 glosses refer to the direct translation of the target vocabulary in the learners' first language, and L2 glosses are the synonyms or short definition of the target word in easier words (Ko, 2012). Previous studies have provided evidence on the effectiveness of both L1 and L2 glosses compared to non-glossed L2 subtitles, but while some studies did not find any significant difference between these two

groups (Jacobs et. al., 1994; Ko, 2012; Yoshii, 2006), others found that L1 glosses differ from L2 glosses (Ertürk, 2016; Laufer & Hill, 2000). For example, Ko (2012) recruited 90 Korean university students who are English learners. The participants were randomly assigned to three conditions: no gloss, L1 gloss, and L2 gloss, and read the materials in their conditions. They took vocabulary tests immediately and four weeks after the intervention. The results showed that both glossed conditions significantly outperformed the no-gloss condition, but there was no significant difference between L1 and L2 glosses both in the immediate and delayed post-tests. However, when this study was replicated by Ertürk (2016), the results were different. Ertürk (2016) recruited 126 preparatory school students in Turkey who are elementary-level learners of English. The results of this study showed that participants who read the materials with L1 glosses significantly outperformed those with L2 glosses in the word form and meaning tests. According to the different results of two same studies, again it could be the case that participants' English proficiency level may influence the effects of different types of glosses. When learners have lower level of proficiency, they may not understand the L2 definitions or synonyms or need longer time to process the glosses. This explanation is also supported by the results of other studies: studies that found no significant difference between L1 and L2 glosses all recruited participants of at least intermediate level (Jachobs et. al., 1994; Ko, 2012; Yoshii, 2006), which is higher than the proficiency of participants in Ertürk's (2016) study. What's more, Laufer & Hill (2000) had participants from high proficiency levels, and found participants with L2 glosses significantly outperformed those with L1 glosses. These results suggest that L2 glosses may be more effective for higher-proficiency learners, and L1 glosses benefit participants with lower proficiency levels more.

Glosses are also used in vocabulary learning through watching videos, but the number of empirical studies that investigated glosses in videos is very small. Montero Perez et. al. (2018) did a study with 227 Dutch-speaking university students who speak

French as L2 and investigated the influence of test announcement (informing the students that a vocabulary test will follow the video or not) and four types of L2 L2 subtitles (no subtitle, full L2 subtitles, L2 keyword subtitles, and glossed L2 keyword subtitles) on L2 vocabulary learning. It was found that students in the glossed L2 keyword subtitles group performed best in both form recognition and meaning recall tests. Teng (2019a) added glossed full L2 subtitles and investigated the effect of four different types of L2 subtitles (glossed full L2 subtitles, glossed L2 keyword subtitles, full L2 subtitles and L2 keyword subtitles) alongside an advance-organizer strategy (present and absent). He recruited 240 Chinese ESL primary school students and randomly assigned them into eight conditions. The findings suggested that glossed full L2 subtitles were the most effective L2 subtitle type. The results of these two studies indicate that L1 glosses in L2 subtitled videos significantly enhanced L2 vocabulary learning for learners with both high and low proficiency levels from drawing their attention to target words and provide meaning of target words directly.

Although both studies provided evidence to support the effectiveness of glosses in L2 subtitled videos, more evidence featuring different types of videos and participants with different proficiency levels and first languages is still needed to support the benefits glosses in L2 subtitled videos have for L2 vocabulary learning. What's more, there are two gaps in the use of glosses that both studies haven't addressed. First, both studies used L1 glosses, which is the direct translation of the L2 target word. However, as reviewed in the studies concerning glosses in written texts, L1 and L2 glosses may have different effects on learners in different proficiency levels (Ertürk, 2016; Laufer & Hill, 2000). Another issue is that the glosses provided in these two studies were not automatic. That is to say, the participants will only access the glosses when they pause the video and click on the target words. Instead of being directly shown as the glosses in written input do, glosses in these videos require an additional "look-up behaviour" (Montero Perez et. al., 2018: 6), just like using a dictionary. This way could be flawed, as it may interrupt the learners' watching process, just like

when they are pausing reading texts to look up for words in a dictionary. What's worse, the participants could just ignore the unfamiliar words and infer the wrong meaning from the contexts (Hulstijn et. al., 1996). Therefore, it could be a better choice to automatically present the glosses when the target words appear in the video.

2.6 The Present Study

The literature review identified several research gaps that few studies have investigated concerning L2 vocabulary learning from watching subtitled videos. First, few studies have compared the effect of L1 subtitles, L2 subtitles, and dual subtitles on L2 vocabulary learning, and the existing studies had mixed results. Besides, no study concerning the comparison between the three types of subtitles was done on young learners with low proficiency levels. Second, few studies have investigated the effectiveness of glosses in L2 subtitled videos. No study has examined the difference between L1 and L2 glosses. Furthermore, previous studies used look-up glosses instead of automatic glosses.

In order to address these gaps this study will explore five different types of subtitling: L2 subtitles, L1 subtitles, dual subtitles, L2 subtitles + L2 glosses, and L2 subtitles + L1 glosses. Besides, this study will use automatic glosses. The present study will recruit Chinese ESL middle school students in Grade 7 and 8. The participants are beginner learners in English and have learnt the language for 5 to 8 years. A general test on vocabulary sizes is used in this study and no pre-tests of the target words to avoid practicing effect and to prevent the learners from paying attention to these words when watching the videos and learning them intentionally. The results of the general vocabulary test are used as covariates to get rid of possible differences in the participants' proficiency levels. A word form test and a word meaning test are used as post-tests to assess the learners' vocabulary gains.

The following research questions are asked in this study:

1. Do subtitles in different languages have different effects on learning L2 new words through watching videos?
2. Do different types of glosses in L2 subtitles have different effects on learning L2 new words through watching subtitled videos?

Chapter 3 Methodology

3.1 Introduction

This chapter provides a detailed introduction to the research methods in this study. First, the participants will be introduced, including their demographic information, the sample size, and the recruitment process. Then, I will introduce the research design, which mainly concerns about how different conditions are decided and how participants are divided into these conditions. After that, research materials and measurements will be introduced. What's more, information on the research procedure, exclusion criteria, and ethical approval will also be given.

3.2 Participants

3.2.1 Sample Size Estimation

Before recruiting the participants, a power analysis with the software G*Power (Faul et. al., 2007) was conducted to estimate the smallest required sample size. Using the ANOVA section in the software, it was calculated that the ideal minimum sample size of this study with a medium effect size is 196 (effect size Cohen's $f = 0.25$, a-level $p = 0.05$, power = 80%, Numerator $df = 4$, number of groups = 5). If power is aimed at 95%, the ideal sample size would be 302. However, as the intervention was conducted online, I aimed at a larger sample size so that participants who did not treat the questionnaires seriously or were not proficient enough to understand the contents of the video could be excluded from the final valid sample.

3.2.2 Recruitment Process

The participants in this study were Year 7-8 Chinese ESL middle school students recruited via the middle school teachers. The teachers were contacted through personal connections. Around 10 schools, 20 classes and 1000 students were invited

to participate the intervention in total so that the sample size may meet the expectation even with a low response rate. The recruitment process of all schools was done all at once. The teachers helped distribute the intervention project to the students' parents who can fill opt-out forms to be excluded from the study. The students who did not return an opt-out form were asked to finish the project as homework at home. The data collection process lasted one week. All incomplete responses were not taken into account in the final sample.

To make the sample more representative, participants from four diverse schools were recruited. School A (55 participants) is located in Nanjing, a major city in Eastern China, and it is one of the best middle schools in China. Students from school A have often received intense English training from early childhood and have a much higher average English proficiency than their peers. School B (30 participants) and School C (11 participants) are average middle schools in Nanjing. They are of similar proficiencies as the pilot participants. School D (14 participants), school E (163 participants) and school F (41 participants) are relatively underachieving. They are located in small towns in Jiangsu Province, and the proficiency of the students is much lower than that of students in big cities. This is because the quality of teachers is lower, and that parents pay less attention to education.

3.2.3 Demographic information

For piloting, 22 participants (12 girls, 10 boys) were recruited. They are Year 7 and 8 middle school students aged between 12 and 14 (M = 12.82, SD = 0.66). All of them are born and raised in China and speak Mandarin Chinese as their first language and English as their second language. They have started learning English as a national curriculum from 6-7 years old. All participants live and study in Nanjing, a major city in eastern China, and are from a school which is above average.

A total of 388 participants finished the formal intervention. After excluding the

invalid data (for exclusion criteria, see 3.7), the data of 314 participants (179 girls, 135 boys) were used for analysis. Even after excluding the invalid responses, the sample size is still above the ideal 302 participants to reach a power of 95%. The participants are Year 7 and 8 students from six middle schools with various socio-economic backgrounds. Their age ranges from 12 to 15 (*M* = 13.54, *SD* = 0.59, *Max* = 15:3, *Min* = 12:10), and they are also Chinese ESL learners, the same as the pilot participants.

3.3 Research Design

This study adopts a quantitative approach and an experimental between-subjects design. There are five conditions to which the participants were randomly assigned to. There is an L2 subtitles group as the baseline condition, in which the videos only include subtitles that accurately reflect the original English contents of the video. Based on this condition, there are two different conditions with different types of subtitles: one with only L1 (Chinese) subtitles and the other with L1 + L2 (Chinese + English) subtitles. There are also two conditions featuring different type of glosses added to the L2 subtitle, one with L1 (Chinese) glosses and the other with L2 (English) glosses. The pilot students tested the materials as they are in the L2 (English) subtitle condition. From establishing the five conditions, this study is able to make two comparisons: the comparison between different types of subtitles and the comparison between different types of glosses. The results of word form and word meaning post-tests serve as dependent variables, and the participants' age and the results of the vocabulary level pre-test serve as control variables.

Table 1 The Five Conditions in this Study.

Condition	Subtitle	Gloss	Number of Participants
1	L2 (English)	N/A	60

2	L1 (Chinese)	N/A	70
3	L1+L2 (English + Chinese)	N/A	64
4	L2 (English)	L2 (English)	59
5	L2 (English)	L1 (Chinese)	61

3.4 Materials

3.4.1 Videos

Three English videos lasting approximately 5 minutes were chosen from YouTube, with two story-telling videos for children and an educational video which delivers knowledge on Peking Opera. All three videos were used in piloting, and one story-telling video and one educational video were chosen as their difficulties fit better with the proficiency of the participants. The difficulty of the words and the speed of the videos are appropriate: over 95% of the words in the videos belong to the 0-2000-word level in the British National Corpus, which allows the participants to easily understand the content of the videos, and therefore concentrate more on the new words while watching (van Zeeland & Schmitt, 2013). On the other hand, the videos are delivered at around 120 words per minute, which is considered as a safe rate for young learners (Tyler et. al., 2009). What's more, during piloting, the students were asked to give their thoughts on the difficulty (easy - difficult) and the speed (fast – slow) of the videos in a 7-point Likert scale. According to the results, both videos have an adequate difficulty ($M1 = 4.09$, $SD1 = 1.19$; $M2 = 4.41$, $SD2 = 0.91$) and speed ($M1 = 4.05$, $SD1 = 0.38$; $M2 = 3.68$, $SD2 = 0.48$).

For each video, all words which are above the 2000-word level in British National Corpus were listed to be potential target words in the intervention, and there were around 30 to 40 words that are listed in each video. Besides, due to the discrepancy

between the frequency level of British National Corpus and Chinese students' learning process, local teachers lent support to decide which words to choose. Five Chinese middle school teachers who have more than ten years' experience in teaching English chose independently the words they think should be new to the students. Following their suggestions, 15 words were chosen for each video. The pilot students were pre-tested with the words, and words that are likely to have been known by the learners were excluded from the list. After combining the results of piloting, suggestions from teachers and the words' distribution in the videos, 10 words were picked for the two videos chosen in the main intervention. For example, the word "gallop", which means running very fast, was chosen as a target word (For the full word list, see Appendix H).

3.4.2 Subtitles and Glosses

Although the original videos have English subtitles embedded in the videos, they are not identical in style and font, and there are some minor mistakes in them. Therefore, all subtitles and glosses were completely remade, and they covered the original subtitles. The Chinese subtitles of the English transcript were translated by the researcher. The translation was done with a literal tendency because the sequence of words in Chinese and English sentences can be largely different if translated freely and liberally. Therefore, in order that the participants could see the exact translation of the voice in the video, the translation delivered the exact sentence-by-sentence meaning, with efforts within each piece of subtitle to make the sentence as natural as possible. Three Chinese middle school teachers with more than 10 years' teaching experience in English gave suggestions on the translation.

The glosses were presented on the right side of the video so that it did not interfere with the video contents. The English glosses were simplified from the definition given by Oxford English Dictionary so that the interpretation is simple, accurate, and easy for learners to comprehend. The Chinese glosses adopted the translations from

the New Oxford English-Chinese Dictionary. Each piece of gloss appeared as soon as the sentence which contains the target word starts and ended after three seconds or when the sentence ends, whichever is later. This is to make sure that the participants have enough time to notice and read the glosses. Figure 3 presents an example of a video with L2 subtitles and L1 glosses.

Figure 3 An example of L2 subtitles + L1 glosses.

3.5 Measures

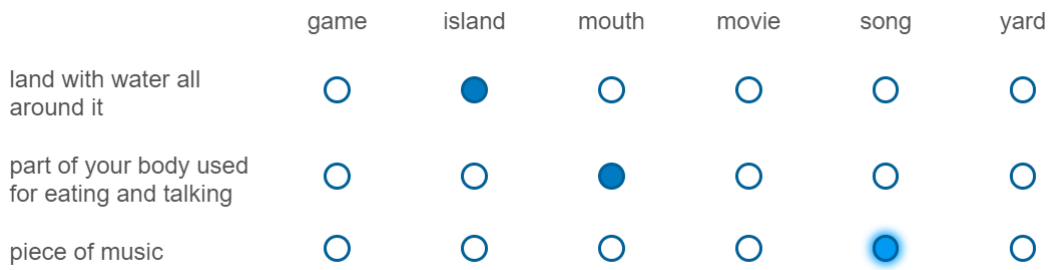
3.5.1 Vocabulary Levels Test

In order to measure the participants' vocabulary size before learning the new words, a pre-test measuring the vocabulary levels was executed. The pre-test did not test the participants' knowledge of the target words, because it may have a practicing effect and draw the learners' attention to the target words in the videos, both of which would influence the accuracy of the post-tests. Webb et. al.'s (2017) updated version of Vocabulary Levels Test was adopted. Vocabulary Levels Test is a test that reflects the students' word frequency level and therefore helps them select which words to learn. It was developed by Nation (1983), and Schmitt et. al., (2001) revised, expanded, and

validated the test. Their tests aim at testing the students' recognition of the form-meaning connections of words at four frequency levels: 2000, 3000, 5000, and 10000. The updated version of Webb et. al. (2017) was used because their new form consists of a more detailed division of levels: 1000, 2000, 3000, 4000, and 5000. As the participants are middle school students whose expected vocabulary size is between 1000 and 2000, the Webb et. al.'s (2017) version is more suitable for this intervention.

The Vocabulary Levels Test adopts a matching format, and there are 30 questions in each level. The maximum score for each level is 30. The words are presented in the format of 10 clusters. For each cluster, there are three definitions, as well as six words to choose from. Three of the words are the correct answers, and the other three words are distractors. The participants are required to choose the correct word for their corresponding definitions. For each 10 clusters, 5 of them are nouns, 3 of them are verbs, and the other 2 adjectives. Figure 4 showed an example of the test (For the full test used in formal intervention, see Appendix E).

Figure 4 An Example of Vocabulary Levels Test.



For piloting, all five levels were administered to see the approximate level of Year 7 and 8 students. The 4th and 5th levels were excluded from the main intervention to increase completion rates, and because no participants reached proficiency in these levels in the pilot. The tests showed high reliability for every level (Level 1 Cronbach's α = .89, Level 2 Cronbach's α = .93, Level 3 Cronbach's α = .93, N = 314).

3.5.2 Comprehension Test

To test whether the participants have watched the video carefully, a simple comprehension test was added after each video. This test wass designed to check whether the participants have paid enough attention to the video and did not serve as a learning outcome measurement. The questions are multiple choice questions with four response options based on the facts in the videos. There are 8 items for piloting, and 5 easiest questions were chosen for each video in the grouped intervention. In piloting, the correct rates of all five chosen items are above 80%, and the chosen questions were made even easier by making the filler options obviously wrong to try to ensure that anyone who has watched the videos would correctly answer all the questions (For the full comprehension test, see Appendix F).

3.5.3 Word Form Test

The word form test consists of multiple choice questions with four response options. For each target word (10 for grouped intervention), there is a question showing the target word and three distracters. The distracters are chosen from words that are similar to the target word in form and those that are synonyms of the target word. For example, for the word “intimidating”, the distracters “imitating” and “intervening” were chosen because they are similar in form, and the distracter “frightening” was also chosen for it has similar meaning to “intimidating”. Participants were required to choose the word they have seen in the video. In piloting, it was found that the correct rate of the test was relatively low, and that it seems that the participants choose the highest frequency word when they were randomly guessing. In some cases, more than half of the participants chose the same word which is not the correct answer. Therefore, in grouped intervention, the distractors were made less similar in form to the original word to make it easier, and the distracters that belong to lower frequency levels than that of the target words are changed. The full word form test can be found

in Appendix G. The reliability for this test is acceptable (Cronbach's $\alpha = .68$).

3.5.4 Word Meaning Test

The word meaning test is the same in format as the word form test. The participants were given the target word (10 for grouped intervention) as well as four Chinese meanings to choose from one of which was a translation of the target meaning. The three distracters belong to the same word type as the target word. For example, for the word "inevitable", the three distracters were "可以食用的" (edible), "神奇的，伟大的" (incredible), and "恐怖的" (horrible). The full word meaning test can be found in Appendix H. The Cronbach's alpha of the word meaning test is .81, indicating good reliability.

3.5.5 Participant Feedback

After each video, participants were also asked to provide feedbacks towards the video and the tests. There are four 7-point Likert scale items for each video, inquiring the participants' subjective feelings towards the difficulty of the video, the difficulty of the comprehension test, the difficulty of the vocabulary test, and the speed of the video. Besides, they are provided with an optional word entry to express their any thoughts towards the contents of the intervention. At the end of the intervention, the participants were given an optional word entry to express their overall feelings towards the intervention. For the full questionnaire, see Appendix I.

3.6 Procedure

The intervention was administered via Qualtrics (https://www.qualtrics.com), an online survey platform which can embed third-party videos. The videos were uploaded to Bilibili (https://www.bilibili.com), a Chinese user-generated content

video platform, and the videos were directly embedded in the Qualtrics project so that the participants will not be redirected to a third-party website. The pilot took around 60-100 minutes to complete, while the formal intervention took approximately 30-50 minutes. Participants started from reading the information sheet and gave assent to participate. Then, they provided their personal information, including name, gender, age, year, and school. Participants first completed the Vocabulary Levels test. Then, after being randomly assigned to one of the five conditions (all pilot participants watched videos with L2 subtitles), the participants watched the two videos in a random sequence (3 for piloting). Timers were added to the videos: participants cannot continue until the video is finished, and the page will automatically jump to the next page one minute after the video is finished so that the participants could only watch the video once. After each video, participants first reported whether they had watched the whole video, and then were asked to complete tests on comprehension and vocabulary and feedbacks towards the videos. The comprehension tests came first, and then vocabulary form test, vocabulary meaning test, and feedback questions were presented in sequence. The participants cannot click the "Next" button if any question is left unanswered. At the end of the survey, participants could provide a final comment on the whole intervention, including but not limited to the difficulty of the videos, the subtitles and glosses, and the experiment design.

3.7 Exclusion Criteria

Because the whole intervention process was conducted online, there is a high possibility that the participants do not treat the experiment seriously. They may not pay full attention while watching the video and may just randomly choose the answers. Besides, as the participants are from various backgrounds, there may be students who are not proficient enough to understand the contents of the videos. To deal with all these possibilities that may do harm to the validity of the data, a series of exclusion criteria were applied to rule out the invalid data.

First, the results of the Vocabulary Levels Test were used to test whether the participants' vocabulary levels are too low for the videos and the target words. The required English vocabulary size for Chinese high school entrance exam which takes place in Year 9 is 1680, therefore, the average vocabulary size of a Year 7 or 8 student should be around 1000-2000. The results of piloting are in accord with this theoretical number: 19 out of 22 participants reached proficiency in the first 1000 level, but only 4 are proficient in the second level. What's more, the lowest score in the first level is 21, which means that students should know most of the first 1000 words. Only 1 student reached proficiency (26 out of 30) in the third level. These results show that as expected, most students have a vocabulary size between 1000 and 2000, and these students should be proficient enough to comprehend the contents of the videos. According to the results of piloting, participants who had a score of lower than 10, the cut-off for chance level performance, in all three levels are excluded – they are either not taking the tasks seriously and randomly guessing or not proficient enough to understand the contents of the video. Thirty participants were excluded for not achieving ideal score in the Vocabulary Levels Test.

Besides, the participants who did not pay full attention to the video are also excluded. First, the participants are asked whether they have watched the full video right after each video. The participants who answered "no" in either of the questions are excluded, because it is highly possible that they have missed some of the new words. 19 participants were excluded because they have not watched the complete video. Another way to check whether the participants have paid full attention in watching the videos is the comprehension test, the answer of which should be obvious for anyone who have watched the full video. There are ten questions in total, and any who did not reach 7 out of 10 are considered as not having paid enough attention to the video and therefore excluded. 25 participants were excluded because of low comprehension score.

In summary, through these exclusion criteria, it is made sure that the participants' proficiency is suitable for the intervention, and that they have put full effort in completing the tasks. After administering the exclusion criteria, a total of 314 participants were included in data analysis.

3.8 Ethical Approval

The whole process of participant recruitment and data management strictly adhered to the guidance of University of Oxford Central University Research Ethics Committee. After receiving the ethical approval (reference: CIA-22HT-043), school teachers were contacted via social media and were provided an information sheet which outlined the purpose of the study, the content of the intervention, and privacy protection (see Appendix). Participants received the information from the teacher and were provided with an opt-out consent form (see Appendix) that they can choose to sign if they were unwilling to participate. Participants who were attending were provided with a link to the Qualtrics project and were asked to finish the intervention at home as homework. The data were collected on a university-owned Qualtrics account. All data were stored on a password-protected laptop during data analysis and will be kept on a university-owned cloud drive for at least three years after the completion of the study.

Chapter 4 Results

This chapter will address the research questions through analyzing the quantitative data from the tests. MANCOVA and between-group ANCOVA tests will be conducted to explore whether different conditions have an influence on word form and meaning learning with vocabulary level as a covariate, and whether there is any significant difference between any two of the five groups.

4.1 Testing Assumptions

Histograms and Q-Q plots were used to visually check the normality of the variables. While results of both word form and word meaning tests conforms to normality, the results of the vocabulary levels tests are not normally distributed. Therefore, non-parametric Mann-Whitney U tests were used to test whether the participants are significantly different in their vocabulary levels before taking the intervention. The results showed that none of the two groups have significantly different results of the vocabulary levels test (all $p > .05$). The results were in accord with the assumption that all participants were randomly assigned to one of the conditions and that their baseline vocabulary knowledge was roughly equivalent for each condition. Scatter plots were used to check whether vocabulary level have a linear relationship with word form and meaning post-tests, and the plots showed that they have linear relationships, which met the assumptions for a covariate in ANCOVA and MANCOVA.

4.2 Descriptive Statistics

The descriptive statistics are summarized in Table 2. It can be seen that the use of glosses in subtitled videos yielded higher scores in word form, while in conditions that include subtitles or glosses in the learners' L1 (Chinese), participants achieved

higher scores in word meaning tests than those in conditions that only include their L2 (English). Among the five conditions, the L2 subtitles + L1 glosses group achieved best results in both word form (7.92) and word meaning (12.79) tests.

Table 2 Descriptive Statistics.

Subtitle	Gloss	n	Vocabulary Level			Comprehension		Word Form		Word Meaning	
			M	Median	SD	M	SD	M	SD	M	SD
L2	N/A	60	53.02	48.00	20.22	9.07	.95	6.98	3.78	9.22	4.57
L1	N/A	70	52.97	50.00	20.89	9.36	.90	6.43	3.29	11.2	3.88
L2+L1	N/A	64	49.47	46.50	18.16	9.56	.77	6.69	3.14	10.86	3.52
L2	L2	59	54.83	49.00	19.58	9.14	.92	7.49	3.55	9.76	4.35
L2	L1	61	55.18	50.00	17.68	8.97	1.03	7.92	3.74	12.79	3.90

4.3 General tests

In order to analyze what effects different conditions have on word form and meaning, the correlation between the tests were first checked. It was found that the results of the word form test and the word meaning test are significantly correlated ($r = .52$, $p < .001$), a multivariate analysis was conducted to look at the overall effects of the independent variable.

A one-way MANCOVA showed a significant main effect of different types of subtitles on the vocabulary post-tests, $F(4, 308) = 6.81$, $p < .001$, $\eta_p^2 = .081$. In the univariate tests, the different types of subtitles showed a significant influence on the word meaning test, but its influence on the word form test is non-significant. The total score of the Vocabulary Levels tests as a covariate showed significant influence on both word form and word meaning tests (see Table 3).

Table 3 Univariate MANCOVA Tests.

Test	Source	*df*	*F*	*p*	η_p^2
Word form	Type of Subtitle	4	1.789	.131	.023
	Vocabulary level	1	194.461	< .001*	.387
Word meaning	Type of Subtitle	4	12.018	< .001*	.135
	Vocabulary level	1	229.663	< .001*	.427
	Error	308			

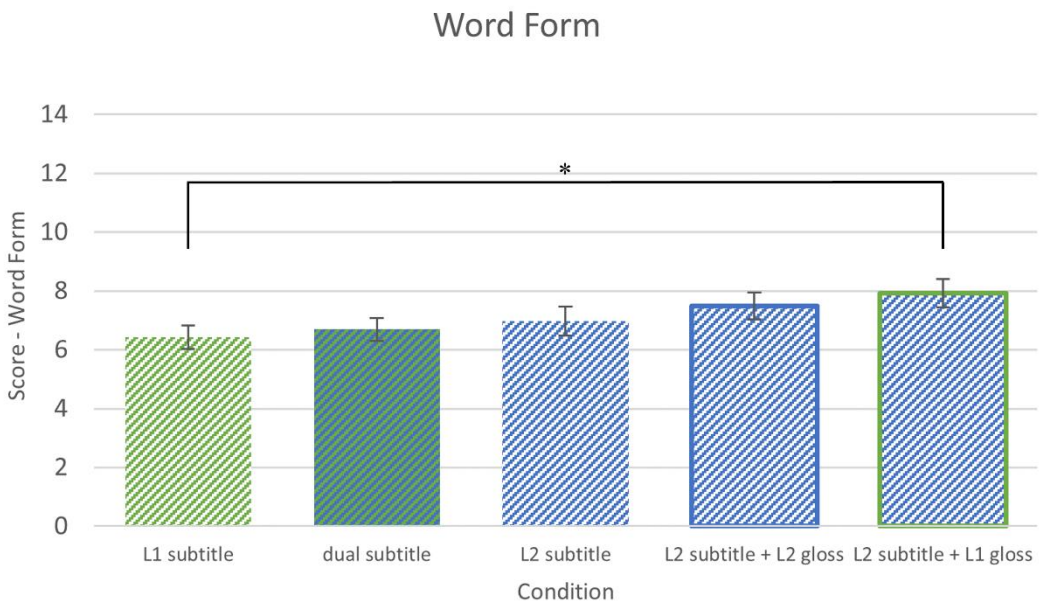
*p < .001.

4.4 Pairwise comparisons

The general MANCOVA was followed by group-to-group pairwise comparisons to see where any significant differences between groups may lie. A total of 10 one-way ANCOVAs were conducted to compare all possible combinations of conditions pairs, with the conditions being the independent variable and vocabulary level as the covariate and either word form or word meaning test as the dependent variable. Table 4 shows the results of the pairwise comparisons among different subtitle types when controlling the pre-test results of the Vocabulary Level test.

For the word form tests, there was only one significant difference between groups: participants in the L2 subtitles + L1 glosses group significantly outperformed participants in the L1 subtitles group. All 9 other comparisons between different conditions did not have a significant result (all p > .05). Figure 5 shows the pairwise comparisons in word form tests.

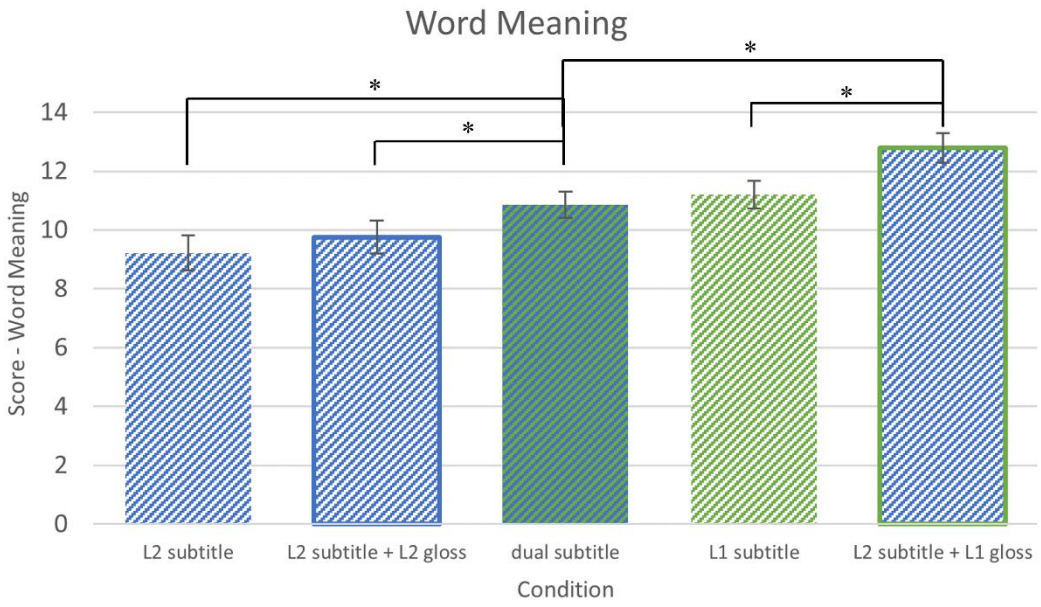
Figure 5 Pairwise Comparisons in Word Form Tests.



* $p < .05$

For the word meaning tests, 8 out of 10 pairs were significantly different, except for the comparison between the L2 subtitles group and the L2 subtitles + L2 glosses group, as well as between the L1 subtitles group and the dual subtitles group. The group with L2 subtitles and L1 glosses was the most effective condition. Participants in this group significantly outperformed participants in every other condition. Figure 6 shows the pairwise comparisons in word form tests.

Figure 6 Pairwise Comparisons in Word Meaning Tests.



* $p < .05$

Table 4 Results of Pairwise Comparisons.

Dependent	Condition		F	p	η_p^2
Form	L2 subtitle	L1 subtitle	1.258	.264	.010
		Dual subtitle	.067	.797	.001
		L2 subtitle + L2 gloss	.321	.572	.003
		L2 subtitle + L1 gloss	1.597	.209	.013
	L1 subtitle	Dual subtitle	1.754	.188	.013
		L2 subtitle + L2 gloss	3.203	.076	.025
		L2 subtitle + L1 gloss	6.415	.013*	.048
	Dual subtitle	L2 subtitle + L2 gloss	.209	.649	.002
		L2 subtitle + L1 gloss	1.199	.276	.010
	L2 subtitle + L2 gloss	L2 subtitle + L1 gloss	.544	.462	.005
Meaning	L2 subtitle	L1 subtitle	11.688	.001**	.084
		Dual subtitle	17.140	< .001***	.124
		L2 subtitle + L2 gloss	.198	.657	.002
		L2 subtitle + L1 gloss	32.030	< .001***	.213

L1 subtitle	Dual subtitle	.019	.892	.000
	L2 subtitle + L2 gloss	8.450	.004**	.063
	L2 subtitle + L1 gloss	5.747	.018*	.043
Dual subtitle	L2 subtitle + L2 gloss	13.196	< .001***	.099
	L2 subtitle + L1 gloss	5.144	.025*	.040
L2 subtitle + L2 gloss	L2 subtitle + L1 gloss	27.845	< .001***	.192

* $p < .05$

** $p < .005$

*** $p < .001$

To clarify where differences lay for the word form and word meaning tests, an exploratory analysis with comparisons involving multiple groups with a certain feature were also conducted. Whether there are glosses, whether there are L1 (Chinese) cues, and whether there are L2 (English) cues are two important distinctions between the groups: two groups have glosses, three groups have L1 cues, and four groups have L2 cues. Two ANCOVAS were used to check whether these two variables have a significant influence on the test results. It was found that groups with L1 cues (L1 subtitle, dual subtitle, and L2 subtitle + L1 gloss) significantly outperformed the groups with only English in the word meaning tests ($F = 40.626$, $p < .001$, $\eta_p^2 = .116$). Besides, groups with glosses significantly outperformed groups without glosses in the word form tests ($F = 4.361$, $p < .05$, $\eta_p^2 = .014$). What's more, the four groups with L2 cues (L2 subtitle, dual subtitle, L2 subtitle + L2 gloss and L2 subtitle + L1 gloss) significantly outperformed the L1 subtitles group which had no L2 cues in the word form tests ($F = 4.908$, $p < .05$, $\eta_p^2 = .016$). None of the comparisons between the remaining tests showed significance.

Chapter 5 Discussion

The present study investigated (1) the impact of subtitles in different languages and (2) the influence of L2 subtitles with glosses in different languages on L2 vocabulary learning through L1 audio-visual input. The study included five treatment groups, and each group had around 60 participants with similar experience in learning the second language who were introduced to 20 new words in two L2 videos. Learners took unexpected tests on word form and meaning knowledge. The findings indicated that subtitling type had a significant effect on vocabulary learning. Among the five groups, using L2 subtitled videos with L1 glosses yielded the best results in both word form and meaning tests. The following parts of this chapter will discuss the results in detail, answer the research questions, articulate the limitations of the study, and list potential directions of future studies.

5.1 Subtitles in different languages

5.1.1 Word form learning with subtitles in different languages

To answer Research Question 1, the results revealed that subtitles in different languages had different contributions to L2 vocabulary form and meaning learning through watching videos. First, the results suggested no significant difference between the three subtitle types in the word form test. The results are in accord with Matielo et. al.'s (2017) and Pujadas & Muñoz's (2019) studies which did not find any significant differences between L2 subtitles and L1 subtitles in the word form tests. However, they are inconsistent with previous studies which found that L2 subtitles are more beneficial than L1 subtitles in word form acquisition (Peters, 2019; Peters et. al., 2016), or Dizon & Thanyawatpokin's (2021) study which found that dual subtitles and L1 subtitles outperformed L2 subtitles.

Despite the overall non-significant result, the trend of the results was such that the L2 subtitles group was slightly higher than the dual subtitles group, followed by L1 subtitles group. The trend for higher results of L2 subtitles than L1 subtitles could indicate that learners may acquire some word forms when provided with the audio-visual input and textual input at the same time. According to Danan's (2004) theory, L2 subtitles may be better suited for intermediate and advanced learners, while L1 subtitles may be more beneficial for beginner learners. Results from previous studies have supported this claim. The studies which found L2 subtitles are better than L1 subtitles in word form learning all had participants in intermediate levels (Peters, 2019; Peters et. al., 2016). By contrast, Dizon & Thanyawatpokin's (2021) study on beginner learners found L1 subtitles were more beneficial than L2 subtitles, and Pujadas & Muñoz's (2019) study on low-intermediate learners found no significant differences between the two types of subtitles. The present study also had learners with relatively low proficiency level. It is possible that these beginners cannot effectively acquire the word forms from the L2 subtitles as they have to put more attention on comprehending the contents of the videos. Such trend of the results is consistent with the claim that L2 subtitles are better for high-proficiency learners. However, the difference is not significant, and more research is needed to confirm the claim.

An alternative explanation for why results in the present study differed from previous studies could be differences in methodology, especially the tests. For those studies which found L2 subtitles were more beneficial, word recognition tests were used, in which participants were provided with target words and are asked whether or not they had heard or read the words in the videos. As it is a self-reported question, it is highly possible that the test results cannot fully represent whether the participants have acquired the word forms: as many tests did not include any "catch" trials in which non-target words are included, this test may depend heavily on honesty from participants, and is likely to be affected by social desirability bias. By contrast, in this

study, multiple choice questions were used, with one target word and three distracters for the participants to choose from. According to the test results, the participants only answered 6 to 7 questions out of 20 correctly on average, which is only a bit higher than randomly guessing (5 out of 20). One possible explanation is that it may be more difficult for the participants to choose the target word from four words that are similar in form or meaning. When the tests are too difficult for the learners, the tests may not be able to discriminate participants in different conditions because all of them may have low scores regardless of the conditions. Another reason could be that the participants of this study only watched each video once, which makes it even harder for them to remember the word forms as they have to put more efforts on comprehending the videos. If they watched the video twice, they could have allocated more attention to the target words, acquired more word forms, and performed better in the word form tests (Montero Perez et. al., 2015).

5.1.2 Word meaning learning with subtitles in different languages

Regarding the results of the word meaning tests, it was found that both the L1 subtitles group and the dual subtitles group significantly outperformed the L2 subtitles group, and that there was no significant difference between the L1 subtitles group and the dual subtitles group. The results accord with the results of the only study that compared the effects of L1 subtitles, L2 subtitles, and dual subtitles (Dizon & Thanyawatpokin, 2021), which also found that both L1 subtitles and dual subtitles have significantly better effects than L2 subtitles in word meaning learning. However, it contradicts with former studies that found L2 subtitles are more beneficial than L1 subtitles in word meaning tests (Peters, 2019; Peters et. al., 2016), and those who did not find any difference between these two types of subtitles (Matielo et. al., 2017; Pujadas & Muñoz, 2019).

One possible explanation for the result is that some learners may fail to derive the meaning of unfamiliar words from L2 subtitles only. Learning the meaning of a new

word is one of the most challenging parts of learning vocabulary (Hasan, 2000; van Zeeland & Schmitt, 2013). This problem may be more severe if the learners are beginners who already faces challenges when trying to comprehend a video in their second language. It may be extremely difficult for them to allocate some of their attention to the new words, and even harder to infer meaning of the unfamiliar words from the contexts. Therefore, when L1 aids are absent, the learners may only learn very few new words. On the contrary, when given the L1 translations of the unfamiliar words, there is a higher possibility that learners acquire initial form-meaning links.

Another explanation is the same as that in the word form section: the proficiency of the learners. As the participants are beginner learners, they may benefit more from L1 subtitles than from L2 subtitles. Beginners may be unable to comprehend the contents of the videos and infer word meaning from contexts when L1 support is absent. In contrast, L1 subtitles could not only help the learners comprehend the videos and allocate more attention to unknown words, but at the same time provide L1 meaning of the unfamiliar words as well. If this explanation were true, it would indicate that the results may only generalize to younger or beginner learners.

From the results, it could be found that the L1 subtitles group and dual subtitles group are not significantly different in both word form and word meaning tests. This indicates that the L2 subtitles may be not playing any role in vocabulary acquisition through dual-subtitled videos. This may be because when provided subtitles in both L1 and L2, the participants could pay attention to only the L1 subtitles which are easier to read and understand. This hypothesis explains why adding L2 subtitles neither facilitates nor hinders the effect of L1 subtitles. Another explanation is that as L2 subtitles had no significant differences from L1 subtitles in word form, the combination of these two could also have similar effects. Regarding of word meaning tests, it could be the case that providing both L1 and L2 subtitles without salience of

target words did not yield more learning of form-meaning links.

5.2 Glosses in different languages

5.2.1 Glosses in word form learning

To answer Research Question 2, the results revealed that adding glosses in different languages to L2 subtitled videos have different effects on vocabulary learning, and that glossed L2 subtitled videos have different effects on vocabulary learning from non-glossed L2 subtitled videos. Regarding the results of word form tests, it was found that among all comparisons, only L1 glosses group significantly outperformed L1 subtitles group. It is worth mentioning that L2 glosses group also showed a trend towards outperforming the L1 subtitles group, though the difference did not reach significance (p = .076). Besides, although not significant, the L1 glosses group and the L2 glosses group have higher scores in word form tests than the three non-glossed conditions. The results could suggest that glosses did enhance L2 vocabulary form learning, but the effect is not large enough to reach a significant difference. The effect of the L1 gloss, and trend for an effect of the L2 gloss on vocabulary form learning is consistent with previous studies on glossed L2 subtitles (Montero Perez et. al., 2018; Teng, 2020), which found that L2 subtitles with L1 glosses resulted in better outcomes in word form learning than non-glossed L2 subtitles. The significant difference between the L1 subtitles group and the L1 glosses group signifies the effectiveness of salient L2 subtitles. Besides, such trend is also supported by the explorative analyses. First, the analysis on L2 cues found that the four groups with L2 cues (L2 subtitles, dual subtitles, L2 subtitles + L1 glosses, L2 subtitles + L2 glosses) significantly outperformed the L1 subtitles group which had no L2 cues in word form tests. The finding suggests that L2 subtitles may be beneficial for the learners' word form learning as they directly provide the forms of the target words. Second, the analysis on glosses revealed that the two groups with glosses significantly

outperformed the three groups without glosses in word form tests. This suggest that glosses, which provide salience in subtitles, could be beneficial for word form learning as they draw the learners' attention to the target words, which may facilitate processing and memorization of target words. The attention-directing cues provided by glosses could draw learners' attention to the target words and therefore provide learners with more opportunities to learn the vocabulary (Neuman et. al., 2018). By contrast, in the non-glossed conditions, the participants may not know which words are the target words.

What's more, there was no significant difference between L1 and L2 glosses for word form learning. It is not surprising because both groups had the same visualisation of target word forms from L2 subtitles, and the difference only lies in the presentation of word meaning. This may suggest that different attention-directing cues do not detract the learners' attention from word form learning. According to previous literature, other ways that provide visual salience of the target words also yielded better word form learning than the ordinary L2 subtitles, including full L2 subtitles with highlighted keywords (Montero Perez et. al., 2014), glossed keyword subtitles (Montero Perez et. al., 2018), and full L2 subtitles with multiple encounters of target words (Teng, 2019a). From the results of these studies as well as the present study, it could be hypothesized that these techniques that provide salience for the target words could similarly draw learners' attention to the words, which may scaffold word form acquisition, and that as long as salience is provided, it does not matter which format or which language the subtitles are in.

5.2.2 Glosses in word meaning learning

The results turned out quite different in the word meaning tests. It was found that similar to previous studies (Montero Perez et. al., 2018; Teng, 2020), L1 glosses had significantly more beneficial effects on learners' word meaning acquisition than L2 subtitles alone. However, L2 glosses, whose effect had not been examined in previous

studies, had no effect on word meaning acquisition at all.

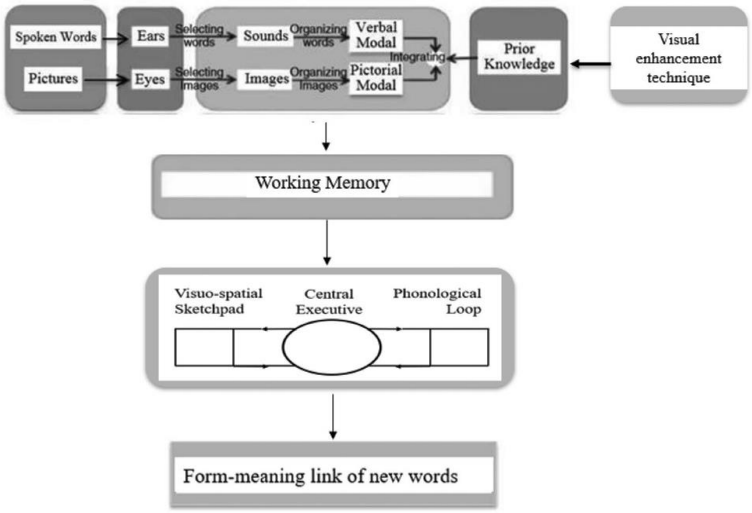
The results suggested that L2 subtitles + L1 glosses group significantly outperformed every other group, including both L2 subtitles group and dual subtitles group. The huge difference between L1 glosses group and non-glossed L2 subtitles group further provides evidence for the claim that inferring word meaning from the context could be a slow and challenging process (Tonzar et. al., 2009), especially for young learners with lower levels of proficiency (Teng, 2019b). When provided with L1 glosses, however, the participants could pay more attention to the target words and directly derive form-meaning links from the glosses.

Although dual subtitles group also have access to the form-meaning link of every word – they even provide more information than L2 subtitles + L1 glosses – they did not enable more word meaning learning compared with L2 subtitles + L1 glosses. Such comparison further supports the importance of salience in on-screen texts. Highlighting the target words with glosses may stimulate their noticing of those words, and the information provided by glosses can further help the learners generate initial form-meaning links.

Teng (2019b, 2020) explained the effectiveness of L2 subtitles and L1 glosses with a dual-model presentation technique (see Figure 7) which is based on Baddeley's (1986, 1992, 2012) working memory model. The effectiveness of L1 glosses could be explained by this theory. It is claimed that L1 glosses may help young learners of a second language manage their working memory more efficiently. With the help of L1 glosses, learners are able to integrate the language input from visuo-spatial sketchpads and phonological loops better. Because L1 glosses provide form-meaning links directly, the videos verbal (aural) and pictorial (visual) information can be more easily integrated. Such integration may better help learners manage their working memory system when incorporating the stimulation from the video and prior

vocabulary knowledge. After that, the working memory systems worked as central executives and managed the two subsystems of storage: visuo-spatial sketchpad and phonological loop. When this process is done, the contents of working memory finally led to a validated form-meaning link of the new words in the storage system.

Figure 7 Teng's (2019b) Model of Subtitled Videos.



By contrast, it is quite surprising that L2 glosses had no effect on word meaning at all. There have been no studies that investigated the effects of L2 glosses in subtitled videos, but the results of this study were consistent with Ertürk's (2016) study which found L1 glosses are more beneficial than L2 glosses for beginners in English in written input. Because results on word form tests suggest that L2 glosses also draws learners' attention to the target word forms, one possible explanation is that the learners failed to acquire form-meaning links from the L2 glosses. One reason that may have contributed to such failure is that the participants of this study are beginners, and the way they have been learning the second language is translating from the first language. That is to say, they are not familiar with learning a new L2 word from L2 interpretation. Another plausible reason is that L2 glosses may be more challenging for the beginners than L1 glosses. In the exploratory analyses, the three groups with L1 cues (L1 subtitles, dual subtitles, L2 subtitles + L1 glosses) significantly outperformed the two groups with only L2 information (L2 subtitles, L2

subtitles + L2 glosses) in the word meaning tests. Although the analysis is explorative in nature and the results cannot be taken as solid evidence, it still shows the trend that L1 cues may be helpful in word meaning learning. The results suggest the importance of L1 information for beginner learners in word meaning learning. The beginner learners may rely heavily on the L1 information, and it could be extremely challenging for them to derive word meaning from contexts when only L2 aids are given. Without L1 cues, comprehending the video could already be challenging for the young learners, and they may not be able to allocate much attention to the meaning of target words, and therefore have generated less form-meaning links.

5.3 Vocabulary size as a covariate

In all 20 ANCOVA tests with word form and word meaning tests as dependent variables respectively, it was found that the vocabulary levels as a covariate had significant and positive effects on vocabulary learning. Learners who scored high on the vocabulary levels test also scored high in the post-tests, indicating the more words a learner knows, the more words they learn. Besides, the effect sizes on all tests are large (all $\eta_p^2 > .339$). The significant relationship and the large effect size suggest that vocabulary size could play an important and major role in vocabulary learning from watching audio-visual input. This finding is consistent with previous studies on vocabulary learning through watching videos with textual aids (Montero Perez et. al., 2018; Peters et. al., 2016).

It should be noted that the present study did not adopt a pre- and post-test design to avoid practicing effect, and post-test results instead of learning gains were used as dependent variables. Therefore, it could be argued that learners may have already known some of the words' meaning and therefore performed better in the word meaning tests. However, learners with larger vocabulary sizes also performed better in the word form test which requires them to recall which word existed in the videos.

Therefore, it could still be argued that learners with larger vocabulary sizes benefited more from the audio-visual input and different types of subtitles.

5.4 Limitations

There are limitations in this study that should be considered when trying to apply these findings. First, the study did not adopt a pre- post-test design. Wanting to make sure that the participants did not expect the vocabulary tests after watching the videos, a pre-test some time before the intervention was not executed. Nor did the study include a control group who were not exposed to the input and only took the vocabulary tests. Therefore, only results in the post-tests instead of learning gains were used in the analyses, and the study cannot give evidence to prove whether learning actually took place in some of the conditions. It is hard to say whether the participants learned those words in the videos, or they had already known the words before the intervention. However, this study included a covariate to control for possible baseline differences in general vocabulary knowledge. Because all groups had no significant differences in general vocabulary knowledge, the findings of group differences should still stand.

Second, no delayed post-tests were executed in this study. Due to restriction of time, only immediate post-tests were administered, and this study therefore could not provide evidence for long-term acquisition. Besides, the study is only a one-time intervention, and no longitudinal effects can be discerned. Hence, this study does not tell us about whether the effects of subtitled videos persist across time or whether there are any differences across different types of subtitles when the intervention is done longitudinally.

Third, the number of encounters with the target words is small. Unlike previous studies which played the video multiple times (Montero Perez et. al., 2014) or make

sure there were at least three encounters of the target words in the video (Teng, 2019c), the present study only presented the short videos once, with most target words used only once or twice. Therefore, there is very limited opportunities for the learners to acquire the target words, which may be different from real-word settings where longer clips with repeated encounters of unfamiliar words are available. However, the videos used in this study were ecologically valid because most videos that learners watch do not deliberately provide repeated encounters of difficult words.

Fourth, the word form test may be too difficult for the learners. Among 20 target words, the learners could only identify around 6-7 target words in the word form tests, which is only a bit higher than random guessing (5 out of 20). Besides, considering the fact that learners have limited encounters with the target words in the videos, the word form test may be even harder for them than if they have multiple exposures. Although the difficulty of word form tests had been lowered after piloting, it is still a bit too hard for the participants. This could be due to the reason that the distractors are very similar. Besides, the reliability of the test is relatively low, indicating that the test still needs improvement. The great difficulty of the form test may be the reason why few significant differences were found between groups.

Fifth, the intervention was done online due to geographical restrictions. Although a series of exclusion criteria were executed to exclude participants who did not watch the videos or answer the questions seriously, it is still possible that some participants were not paying enough attention to the contents of the intervention. What's more, although time restrictors were used to make sure the videos are only watched once, it is still possible that participants paused, winded back, or forwarded the videos. This indicates that participants could have performed better in the tests if the intervention had been done off-line.

Sixth, although the participants are of similar age, their English proficiency levels are

not fully homogeneous. The participants were recruited from several different middle schools, with various socio-economic backgrounds. There are participants from top schools who easily got high scores in all tests, and there are those from schools in underrepresented areas who did not recognize the easiest words from 0-1000 level. Therefore, the audio-visual input may be very easy for some of the participants, but on the other hand may be too hard for other participants to comprehend. In the survey that asked the participants' feelings towards the intervention, a few participants complained that the videos and tests are too difficult for them even after the tests were piloted and made easier. That might be the reason why the result of vocabulary levels test was not normally distributed.

5.5 Future research directions

Considering the findings and limitations of this study, some suggestions can be given for future studies. Future studies should first test different types of subtitles for learners of different age groups and learners in different proficiency levels. It is highly possible that some types of subtitles are more beneficial for advanced learners, and others are more suitable for beginners. Besides, whether the acquisition of target words can be retained in the period of time could be addressed. Apart from that, studies on the long-term learning effect of L2 vocabulary learning through watching subtitled videos are also needed. These studies could provide the participants with videos periodically over several months and test their vocabulary growth through this period of time. What's more, although this study revealed the effects of different types of subtitles, it is still unclear how they function. Eye-tracking studies may be helpful for us to understand whether participants read different types of subtitles for different number of times.

5.6 Pedagogical implications

This study is innovative in investigating subtitles and glosses in different languages. The results of this study provide important pedagogical implications. First, the findings of the present study reveal that access to meaning may be helpful for L2 vocabulary form-meaning links formation through watching audio-visual material, especially for beginner learners. Access to L1 meaning, no matter in the form of L1 subtitles or L1 glosses, could not only facilitate learners' comprehension of the contents of the video, but also yield initial form-meaning link construction. By contrast, materials that are fully in the second language may not be suitable for beginners, even if interpretations of unfamiliar words in L2 are provided, learners may still struggle to understand. Beginners may heavily rely on L1 translations when they are learning new words, and it could be too challenging for them if L1 cues are not offered.

Second, glosses in L2 subtitled videos should be encouraged. It not only draws learners' attention to the target words, but also offers form-meaning links so that the learners could acquire new vocabulary more easily. It promotes students' word form recognition and form-meaning connection at the same time. Therefore, when creating L2 subtitles for a L2 learning video, it is recommended that glosses are provided alongside the L2 subtitles to foster more intentional learning of target words. L1 glosses are more recommended for young learners who are beginner learners.

Third, among all different types of subtitles, the combination of L2 subtitles and L1 glosses is the most effective type for beginner learners. This type of subtitle could provide word forms, stimulate noticing of target words, and provide learners with form-meaning links at the same time, which could benefit both word form and word meaning learning. It is recommended that learners choose videos that provide L2 subtitles with L1 glosses, and that teachers could provide this type of videos in classrooms.

Chapter 6 Conclusion

The study conducted an experiment on beginner learners of English as a second language to investigate the effect of different types of subtitles on L2 vocabulary learning through watching videos. This study adds evidence to the literature about subtitles in different languages and use of glosses in subtitled videos. In terms of subtitles in different languages, the results show that L1 subtitles and dual subtitles had significantly better effect in word meaning learning than L2 subtitles, while these three types of subtitles had no significant difference in their effects on word form learning. In terms of the use of glosses in L2 subtitled videos, it was found that the L1 glosses group significantly outperformed the L1 subtitles group in word form learning, and while the L1 glosses group significantly outperformed every other group in the word meaning tests, the L2 glosses group had no significant difference from L2 subtitles group in word meaning learning. The results of the current study suggest that L1 cues, no matter in the format of subtitles or glosses, may be important for beginner learners to acquire initial form-meaning links because these learners may not be capable of inferring the word meaning from contexts without the support of their first language. Besides, it is suggested that glosses are effective in enhancing word form and meaning learning. Glosses could not only stimulate the learners' noticing toward target words so that they may acquire more knowledge about the word form, but also provide interpretations of the words to scaffold initial acquisition of word form-meaning links. The combination of L2 subtitles and L1 glosses is considered as the most beneficial condition both in word form and word meaning learning and is recommended to beginner learners of a second language. The pedagogical implications of this study will contribute to more effective use of subtitled videos in L2 vocabulary learning, especially in the context of COVID-19 where face-to-face instructions are limited, and online materials are important resources for language learning.

References

Abraham, L. B. (2008). Computer-mediated glosses in second language reading comprehension and vocabulary learning: A meta-analysis. *Computer Assisted Language Learning*, *21*(3), 199–226. https://doi.org/10.1080/09588220802090246

Almeida, P. A., & Costa, P. D. (2014a). Foreign Language Acquisition: The Role of Subtitling. *Procedia - Social and Behavioral Sciences*, *141*, 1234–1238. https://doi.org/10.1016/j.sbspro.2014.05.212

Almeida, P. A., & Costa, P. D. (2014b). Foreign Language Acquisition: The Role of Subtitling. *Procedia - Social and Behavioral Sciences*, *141*, 1234–1238. https://doi.org/10.1016/j.sbspro.2014.05.212

Baddeley, A. (1986). *Working memory* (pp. xi, 289). Clarendon Press/Oxford University Press.

Baddeley, A. (1992). Working memory. *Science*, *255*(5044), 556–559.

Baddeley, A. (2012). Working Memory: Theories, Models, and Controversies. *Annual Review of Psychology*, *63*(1), 1–29. https://doi.org/10.1146/annurev-psych-120710-100422

Baltova, I. (1994). The Impact of Video on the Comprehension Skills of Core French Students. *Canadian Modern Language Review*, *50*(3), 507–531.

Barcroft, J. (2004). Second Language Vocabulary Acquisition: A Lexical Input Processing Approach. *Foreign Language Annals*, *37*(2), 200–208. https://doi.org/10.1111/j.1944-9720.2004.tb02193.x

Barcroft, J. (2009). Strategies and performance in intentional L2 vocabulary learning. *Language Awareness*, *18*(1), 74–89. https://doi.org/10.1080/09658410802557535

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. https://doi.org/10.1017/S0142716402004022

Bisson, M.-J., Heuven, W. J. B. V., Conklin, K., & Tunney, R. J. (2014). Processing of native and foreign language subtitles in films: An eye tracking study. *Applied Psycholinguistics*, *35*(2), 399–418. https://doi.org/10.1017/S0142716412000434

Bobis, J., Sweller, J., & Cooper, M. (1993). Cognitive load effects in a primary-school geometry task. *Learning and Instruction*, *3*(1), 1–21. https://doi.org/10.1016/S0959-4752(09)80002-9

Brett, P. (1995). Multimedia for listening comprehension: The design of a multimedia-based resource for developing listening skills. *System*, *23*(1), 77–85. https://doi.org/10.1016/0346-251X(94)00054-A

Brett, P. (1998). Using Multimedia: A Descriptive Investigation of Incidental Language Learning. *Computer Assisted Language Learning*, *11*(2), 179–200. https://doi.org/10.1076/call.11.2.179.5684

Chai, J., & Erlam, R. (2008). The Effect and the Influence of the Use of Video and Captions on Second Language Learning. *New Zealand Studies in Applied Linguistics*, *14*(2), 25–44. https://doi.org/10.3316/informit.997746379301329

Cintrón-Valentín, M., García-Amaya, L., & Ellis, N. C. (2019). Captioning and grammar learning in the L2 Spanish classroom. *The Language Learning Journal*, *47*(4), 439–459. https://doi.org/10.1080/09571736.2019.1615978

Coady, J., & Huckin, T. (1997). *Second Language Vocabulary Acquisition: A Rationale for Pedagogy*. Cambridge University Press.

d'Ydewalle, G., & Pavakanun, U. (1995). Acquisition of a Second/Foreign Language by Viewing a Television Program. In P. Winterhoff-Spurk (Ed.), *Psychology of Media in Europe: The State of the Art—Perspectives for the Future* (pp. 51–64). VS Verlag für Sozialwissenschaften. https://doi.org/10.1007/978-3-322-94251-7_6

d'Ydewalle, G., & Pavakanun, U. (1997). Could Enjoying a Movie Lead to Language Acquisition? In P. Winterhoff-Spurk & T. H. A. van der Voort (Eds.), *New Horizons in Media Psychology: Research Cooperation and Projects in Europe* (pp. 145–155). VS Verlag für Sozialwissenschaften. https://doi.org/10.1007/978-3-663-10899-3_10

d'Ydewalle, G., & Van de Poel, M. (1999). Incidental Foreign-Language Acquisition by Children Watching Subtitled Television Programs. *Journal of Psycholinguistic Research*, *28*(3), 227–244. https://doi.org/10.1023/A:1023202130625

Danan, M. (2004). Captioning and Subtitling: Undervalued Language Learning Strategies. *Meta : Journal Des Traducteurs / Meta: Translators' Journal*, *49*(1), 67–77. https://doi.org/10.7202/009021ar

Dizon, G., & Thanyawatpokin, B. (2021). *Language learning with Netflix: Exploring the effects of dual subtitles on vocabulary learning and listening comprehension*. 14.

Duquette, L., & Painchaud, G. (1996). A Comparison of Vocabulary Acquisition in Audio and Video Contexts. *The Canadian Modern Language Review*, *53*(1), 143–172. https://doi.org/10.3138/cmlr.53.1.143

Ertürk, Z. ÿdem. (2016). The Effect of Glossing on EFL Learners Incidental Vocabulary Learning in Reading. *Procedia - Social and Behavioral Sciences*, *232*, 373–381. https://doi.org/10.1016/j.sbspro.2016.10.052

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, *39*(2), 175–191. https://doi.org/10.3758/BF03193146

File, K. A., & Adams, R. (2010). Should Vocabulary Instruction Be Integrated or Isolated? *TESOL Quarterly*, *44*(2), 222–249. https://doi.org/10.5054/tq.2010.219943

Fletcher, J. D., & Tobias, S. (2005). The Multimedia Principle. In *The Cambridge handbook of multimedia learning* (pp. 117–133). Cambridge University Press. https://doi.org/10.1017/CBO9780511816819.008

Frumuselu, A. D., De Maeyer, S., Donche, V., & Colon Plana, M. del M. G. (2015). Television series inside the EFL classroom: Bridging the gap between teaching and learning informal language through subtitles. *Linguistics and Education*, *32*, 107–117. https://doi.org/10.1016/j.linged.2015.10.001

Garza, T. J. (1991). Evaluating the Use of Captioned Video Materials in Advanced Foreign Language Learning. *Foreign Language Annals*, *24*(3), 239–258. https://doi.org/10.1111/j.1944-9720.1991.tb00469.x

Gass, S. (1999). DISCUSSION: Incidental Vocabulary Learning. *Studies in Second Language Acquisition*, *21*(2), 319–333. https://doi.org/10.1017/S0272263199002090

Guillory, H. G. (1998). The Effects of Keyword Captions to Authentic French Video on Learner Comprehension. *CALICO Journal*, *15*(1–3), 89–108. https://doi.org/10.1558/cj.v15i1-3.89-108

Hasan, A. S. (2000). Learners' Perceptions of Listening Comprehension Problems. *Language, Culture and Curriculum*, *13*(2), 137–153. https://doi.org/10.1080/07908310008666595

Hong, X. (2010). Review of Effects of Glosses on Incidental Vocabulary Learning and Reaing Comprehension. *Chinese Journal of Applied Linguistics (Foreign Language Teaching & Research Press)*, *33*(1).

Hsieh, Y. (2020). Effects of video captioning on EFL vocabulary learning and listening comprehension. *Computer Assisted Language Learning*, *33*(5–6), 567–589. https://doi.org/10.1080/09588221.2019.1577898

Hulstijn, J. H. (2003). Incidental and Intentional Learning. In *The Handbook of Second Language Acquisition* (pp. 349–381). John Wiley & Sons, Ltd. https://doi.org/10.1002/9780470756492.ch12

Hulstijn, J. H., Hollander, M., & Greidanus, T. (1996). Incidental Vocabulary Learning by Advanced Foreign Language Students: The Influence of Marginal Glosses, Dictionary Use, and Reoccurrence of Unknown Words. *The Modern Language Journal*, *80*(3), 327–339. https://doi.org/10.1111/j.1540-4781.1996.tb01614.x

Jacobs, G. M., Dufon, P., & Hong, F. C. (1994). L1 and L2 vocabulary glosses in L2 reading passages: Their effectiveness for increasing comprehension and vocabulary knowledge. *Journal of Research in Reading*, *17*(1), 19–28. https://doi.org/10.1111/j.1467-9817.1994.tb00049.x

Ko, M. H. (2005). *Glosses, comprehension, and strategy use*. http://hdl.handle.net/10125/66786

Ko, M. H. (2012). Glossing and Second Language Vocabulary Learning. *TESOL Quarterly*, *46*(1), 56–79. https://doi.org/10.1002/tesq.3

Koolstra, C. M., & Beentjes, J. W. J. (1999). Children's vocabulary acquisition in a foreign language through watching subtitled television programs at home. *Educational Technology Research and Development*, *47*(1), 51–60. https://doi.org/10.1007/BF02299476

Laufer, B. (2003). Vocabulary Acquisition in a Second Language: Do Learners Really Acquire Most Vocabulary by Reading? Some Empirical Evidence. *The Canadian Modern Language Review*, *59*(4), 567–587. https://doi.org/10.3138/cmlr.59.4.567

Laufer, B. (2005). Focus on Form in Second Language Vocabulary Learning. *EUROSLA Yearbook*, *5*(1), 223–250. https://doi.org/10.1075/eurosla.5.11lau

Laufer, B. (2010). Chapter 2. Form-focused Instruction in Second Language Vocabulary Learning. In *Chapter 2. Form-focused Instruction in Second Language Vocabulary Learning* (pp. 15–27). Multilingual Matters. https://doi.org/10.21832/9781847692900-003

Laufer, B., & Hill, M. (2000). *What Lexical Information Do L2 Learners Select in a CALL Dictionary and How Does It Affect Word Retention?* For full text: http://llt. https://eric.ed.gov/?id=ED462834

Lomicka, L. (1998). "To Gloss or Not to Gloss": An Investigation of Reading Comprehension Online. *Language Learning & Technology*, *1*(2), 41–50.

Lwo, L., & Lin, M. C.-T. (2012). The effects of captions in teenagers' multimedia L2 learning. *ReCALL*, *24*(2), 188–208. https://doi.org/10.1017/S0958344012000067

Markham, P. (1999). Captioned Videotapes and Second-Language Listening Word Recognition. *Foreign Language Annals*, *32*(3), 321–328. https://doi.org/10.1111/j.1944-9720.1999.tb01344.x

Matielo, R., Oliveira, R. P. de, & Baretta, L. (2017). Intralingual subtitles, interlingual subtitles, and video comprehension: Insights from an exploratory study. *Letrônica*, *10*(2), 758–774. https://doi.org/10.15448/1984-4301.2017.2.26370

Mayer, R. E. (2001). *Multimedia Learning*. Cambridge University Press. https://doi.org/10.1017/CBO9781139164603

Mayer, R., & Mayer, R. E. (2005). *The Cambridge Handbook of Multimedia Learning*. Cambridge University Press.

Meskill, C. (1996). Listening Skills Development Through Multimedia. *Journal of Educational Multimedia and Hypermedia*, *5*(2), 179–201.

Montero Perez, M. (2022). Second or foreign language learning through watching audio-visual input and the role of on-screen text. *Language Teaching*, *55*(2), 163–192. Scopus. https://doi.org/10.1017/S0261444821000501

Montero Perez, M., Peters, E., Clarebout, G., & Desmet, P. (2014). Effects Of Captioning On Video Comprehension And Incidental Vocabulary Learning. *Language Learning & Technology*, *18*(1), 118–141.

Montero Perez, M., Peters, E., & Desmet, P. (2015). Enhancing Vocabulary Learning Through Captioned Video: An Eye-Tracking Study. *The Modern Language Journal*, *99*(2), 308–328. https://doi.org/10.1111/modl.12215

Montero Perez, M., Peters, E., & Desmet, P. (2018). Vocabulary learning through viewing video: The effect of two enhancement techniques. *Computer Assisted Language Learning*, *31*(1–2), 1–26. https://doi.org/10.1080/09588221.2017.1375960

Montero Perez, M., Van Den Noortgate, W., & Desmet, P. (2013). Captioned video for L2 listening and vocabulary learning: A meta-analysis. *System*, *41*(3), 720–739. https://doi.org/10.1016/j.system.2013.07.013

Nagata, N. (1999). The Effectiveness of Computer-Assisted Interactive Glosses. *Foreign Language Annals*, *32*(4), 469–479. https://doi.org/10.1111/j.1944-9720.1999.tb00876.x

Nation, I. (2006). How large a vocabulary is needed for reading and listening? *Canadian Modern Language Review*, *63*(1), 59–82.

Nation, I. S. P. (2001). *Learning Vocabulary in Another Language*. Cambridge University Press. https://doi.org/10.1017/CBO9781139524759

Neuman, S. B., Wong, K. M., Flynn, R., & Kaefer, T. (2019). Learning vocabulary from educational media: The role of pedagogical supports for low-income preschoolers. *Journal of Educational Psychology*, *111*(1), 32–44. https://doi.org/10.1037/edu0000278

Nielsen, R. J. (2016). *Japanese Vocabulary Learning Through an Interactive Video Platform: Comparative Effects of L1 Versus L2 Definitions and Kana Versus Kanji Presentation* [M.A., Brigham Young University]. https://www.proquest.com/docview/2446699921/abstract/3F080115435C431BPQ/1

Paivio, A. (1986). *Mental Representations: A Dual Coding Approach*. Oxford University Press.

Park, M. (2004). *The effects of partial captions on Korean EFL learners' listening comprehension* [Ph.D., The University of Texas at Austin]. https://www.proquest.com/docview/305124957/abstract/7B8B1DE1C99B4C27PQ/1

Peters, E. (2019). The Effect of Imagery and On-Screen Text on Foreign Language Vocabulary Learning from Audiovisual Input. *TESOL Quarterly*, *53*(4), 1008–1032. https://doi.org/10.1002/tesq.531

Peters, E., Heynen, E., & Puimège, E. (2016). Learning vocabulary through audiovisual input: The differential effect of L1 subtitles and captions. *System*, *63*, 134–148. https://doi.org/10.1016/j.system.2016.10.002

Pujadas, G., & Muñoz, C. (2019). Extensive viewing of captioned and subtitled TV series: A study of L2 vocabulary learning by adolescents. *The Language Learning Journal*, *47*(4), 479–496. https://doi.org/10.1080/09571736.2019.1616806

Qian, D. D. (1996). ESL Vocabulary Acquisition: Contextualization and Decontextualization. *The Canadian Modern Language Review*, *53*(1), 120–142. https://doi.org/10.3138/cmlr.53.1.120

Raine, P., SONDA, N., & KRAUSE, A. (2013). Incidental increase in depth of vocabulary knowledge through the viewing of subtitled, authentic videos. *Making a Difference. Tokyo: JALT*, 492–505.

Sadoski, M., Paivio, A., & Goetz, E. T. (1991). Commentary: A Critique of Schema Theory in Reading and a Dual Coding Alternative. *Reading Research Quarterly*, *26*(4), 463. https://doi.org/10.2307/747898

Schmitt, N., Schmitt, D., & Clapham, C. (2001). Developing and exploring the behaviour of two new versions of the Vocabulary Levels Test. *Language Testing*, *18*(1), 55–88. https://doi.org/10.1177/026553220101800103

Sweller, J. (2005). The Redundancy Principle in Multimedia Learning. In R. Mayer (Ed.), *The Cambridge Handbook of Multimedia Learning* (pp. 159–168).

Cambridge University Press. https://doi.org/10.1017/CBO9780511816819.011

Sydorenko, T. (2010). Modality of Input and Vocabulary Acquisition. *Language Learning*, 24.

Tajeddin, Z., & Daraee, D. (2013). Vocabulary Acquisition through Written Input: Effects of Form-Focused, Message-Oriented, and Comprehension Tasks. *TESL-EJ*, *16*(4). https://eric.ed.gov/?id=EJ1004356

Teng, F. (2019a). Incidental vocabulary learning for primary school students: The effects of L2 caption type and word exposure frequency. *The Australian Educational Researcher*, *46*(1), 113–136. https://doi.org/10.1007/s13384-018-0279-6

Teng, F. (2019b). Maximizing the potential of captions for primary school ESL students' comprehension of English-language videos. *Computer Assisted Language Learning*, *32*(7), 665–691. https://doi.org/10.1080/09588221.2018.1532912

Teng, F. (2019c). The effects of video caption types and advance organizers on incidental L2 collocation learning. *Computers & Education*, *142*, 103655. https://doi.org/10.1016/j.compedu.2019.103655

Teng, F. (2020). Vocabulary learning through videos: Captions, advance-organizer strategy, and their combination. *Computer Assisted Language Learning*, *0*(0), 1–33. https://doi.org/10.1080/09588221.2020.1720253

Teng, F. (2021). *Language Learning Through Captioned Videos: Incidental Vocabulary Acquisition*. Routledge.

Teng, F. (2022). Incidental L2 vocabulary learning from viewing captioned videos: Effects of learner-related factors. *System*, *105*, 102736. https://doi.org/10.1016/j.system.2022.102736

Tonzar, C., Lotto, L., & Job, R. (2009). L2 Vocabulary Acquisition in Children: Effects of Learning Method and Cognate Status. *Language Learning*, *59*(3), 623–646. https://doi.org/10.1111/j.1467-9922.2009.00519.x

Tyler, M. D., Jones, C., Grebennikov, L., Leigh, G., Noble, W., & Burnham, D. (2009). Effect of caption rate on the comprehension of educational television programmes by deaf school students. *Deafness and Education International*, *11*(3), 152–162. https://doi.org/10.1002/dei.262

van Zeeland, H., & Schmitt, N. (2013). Incidental vocabulary acquisition through L2 listening: A dimensions approach. *System*, *41*(3), 609–624. https://doi.org/10.1016/j.system.2013.07.012

Vanderplank, R. (2010). Déjà vu? A decade of research on language laboratories, television and video in language learning. *Language Teaching*, *43*(1), 1–37. https://doi.org/10.1017/S0261444809990267

Vulchanova, M., Aurstad, L. M. G., Kvitnes, I. E. N., & Eshuis, H. (2015). As naturalistic as it gets: Subtitles in the English classroom in Norway. *Frontiers in Psychology*, *5*. https://www.frontiersin.org/articles/10.3389/fpsyg.2014.01510

Wang, A., & Pellicer-Sánchez, A. (2022). Incidental Vocabulary Learning From Bilingual Subtitled Viewing: An Eye-Tracking Study. *Language Learning*, *72*(3), 765–805. https://doi.org/10.1111/lang.12495

Wang, Y. T. (2019). Effects of L1/L2 Captioned TV Programs on Students' Vocabulary Learning and Comprehension. *Calico Journal*, *36*(3), 204–224. https://doi.org/10.1558/cj.36268

Webb, S. (2020). Incidental Vocabulary Learning. In *The Routledge Handbook of Vocabulary Studies*. Routledge.

Webb, S., Sasao, Y., & Ballance, O. (2017). The updated Vocabulary Levels Test: Developing and validating two new forms of the VLT. *ITL - International Journal of Applied Linguistics*, *168*(1), 33–69. https://doi.org/10.1075/itl.168.1.02web

Wei, R., & Fan, L. (2022). On-Screen Texts in Audiovisual Input for L2 Vocabulary Learning: A Review. *Frontiers in Psychology*, *13*. Scopus. https://doi.org/10.3389/fpsyg.2022.904523

Winke, P., Gass, S., & Sydorenko, T. (2010). The Effects of Captioning Videos Used for Foreign Language Listening Activities. *Language Learning*, 22.

Winke, P., Gass, S., & Sydorenko, T. (2013). Factors Influencing the Use of Captions by Foreign Language Learners: An Eye-Tracking Study. *The Modern Language Journal*, *97*(1), 254–275. https://doi.org/10.1111/j.1540-4781.2013.01432.x

Wong, K. M., Flynn, R. M., & Neuman, S. B. (2021). L2 vocabulary learning from educational media: The influence of screen-based scaffolds on the incidental–intentional continuum. *TESOL Journal*, *12*(4), e641. https://doi.org/10.1002/tesj.641

Yoshii, M. (2006). L1 and L2 Glosses: Their Effects on Incidental Vocabulary Learning. *Language Learning*, 17.

Appendix A Recruitment Email

Pilot & Intervention

Dear **[teacher name]**
I hope you are well. My name is xxxxxx and I am a Master student at the Department of Education at University of Oxford. I am conducting a research project with Year 7-9 students which we wanted to invite **[name of school]** to participate in.
The goal of the project is to develop materials to help students learn new vocabulary through watching different types of videos. This includes watching English videos and doing some language quizzes using a computer.

The study is for Year 7-9 students and is taking place **[dates]**. In the study, the students will complete an online task which will be set as homework to be done at home. In the online task, the students are expected to fill some demographic information, do some English language tests, watch the given videos and take quizzes about the video. We would give opt-out consent forms to all parents prior to the study, as we'd like to try to ensure that all students have the opportunity to participate. The task will last around 40-50 minutes.

If you're happy, I'd be grateful if you could pass this email onto your Year 7-9 teacher to discuss the details and scheduling.

In return for your help, we would be very happy to share our results and materials with you at the end of the study.

Please let me know if you have any questions or need any further information.

Many thanks,
xxxxxx

Appendix B Information Sheet and Opt-out Form

Dear Parent/Guardian,

My name is Yaguang Li and I am conducting a study about middle school students' English vocabulary learning as part of my Master's dissertation at the University of Oxford. Your school has kindly agreed to take part. Please read and keep this letter to understand what the study is about, why we're doing it and what it would involve for you and your child.

I will be giving online learning activities about English vocabulary to Year 7-9 students which will be set as homework. Students will complete an online task that lasts around 40-50 minutes. During the task, students will do some English language quizzes and watch an English video using a computer. This is so we can see how different learning activities benefit children's English learning. We will also ask them a few questions about themselves, including their age and gender, so we know who was in our sample.

The results of this research will help us to find the most beneficial approach for learning English vocabulary. There are no identified risks to taking part in this research. If your child does not wish to take part, or wishes to stop taking part at any time, they can stop doing the task.

Your child's responses will be confidential and will not be shared with their school or their teachers. A summary of the findings which does not identify individual children's results will be made available to teachers and parents at the end of the research, and findings will be presented in a Master's dissertation, and may be shared at academic conferences or published in academic journals. If you wish to withdraw your child's data, you may do so by emailing the researcher by the end of June 2022, or one month after data collection, whichever is later. Members of our research team only will have access to the research data, which will be stored on a secure cloud-based server and/or in a locked filing cabinet at the University of Oxford. Research data will be kept for a minimum of 3 years after publication of our research findings. The study has been approved by the University of Oxford department of education ethics committee, reference: [CIA-22HT-043].

If you are happy for your child to take part in our research you do not need to take any action.

However, if you **do not wish** for your child to participate, please fill in the form to opt out below.

Feel free to contact me if you have any questions about participating in the study, or to receive a copy of the results when the study is complete. I am grateful for your help with my research.

Yaguang Li

– –

If you DO NOT want your child to participate, please complete and return this slip to school by [date].

I do not want my child to take part in Yaguang Li's vocabulary study.

Child's name: ________________________________

Child's school class: _______________

Parent's signature: __________________________ Date: _____________

Appendix C Assent Script

Assent script for participating students

Dear student, thank you for participating in my research. Today you are going to learn English from watching an English video. There will be several small quizzes before and after watching the English video. If you would like to participate, please click 'Continue' and start the activities.

Survey information

What is this project about?

This survey is being conducted by researchers at the Oxford University Department of Education to investigate how watching an English video may influence English language learning. We plan to use the findings to share ideas that may be useful to parents and schools to support language learning and in academic publications.

Can I take part?

To take part, you must be a Chinese middle school student with English as your second language.

What does the task involve?

The task will take around 50 minutes to complete. You will have to complete it in one sitting. In this task, you will take some quizzes in English and watch English videos. You will also be asked some basic demographic information about yourself so that we know who has completed our survey. If you find yourself uncomfortable when doing the tasks, you can opt not to continue the task if you wish.

Is the data anonymous?

We ask for your name so that you can withdraw your information if you wish. Your name will be deleted from the data right after the deadline of withdrawing your data. Your individual data will NOT be shared with your teacher or parents.

Can I find out the results of the project?

We will send your teacher a brief report summarising the main findings of the survey at the end of the project. You can ask your teacher for the report.

When do I need to complete the survey?

Please complete the survey by [DATE].

Will I be paid to take part?

You will not be paid but we will send small gifts to your class as a whole.

What if I change my mind?

You can withdraw from the study by closing the browser and tell your teacher about it.

Where can I find out more about the study?

If you have any questions about the study, before or after taking part, please tell your teacher or contact Yaguang Li. This research has received ethical approval from the University of Oxford Department of Education ethics committee (reference: CIA-22HT-043)

Appendix D Video Scripts

Video A

David and Goliath

大卫和歌利亚

This is not an ordinary story,

这不是一个普通的故事，

for never has there ever been a story like this:

因为从来不会有像这样的故事：

a true story of contrasts, of **faith**,

一个真实的，充满着矛盾、信仰、

and the spirit of never giving up.

以及永不言弃的精神的故事。

This … is the story of David and Goliath.

这……就是大卫和歌利亚的故事。

And it starts right here on a beautiful sunny day

这个故事开始于一个美丽的晴天，

in the lush green meadow.

发生在苍翠的牧场中。

Dear God, thank you for this beautiful day!

亲爱的神明，感谢你给予我们如此美丽的一天！

This is David, a kind, god-loving and happy shepherd.

这是大卫，一个善良、虔诚和快乐的牧羊人。

David was the youngest of eight brothers,

大卫是家里八个兄弟中最年幼的，

and spent all his day watching his sheep.

他整天都在放羊。

His brothers however were soldiers of their country’s army,

然而，他的哥哥们都是他们国家军队的士兵，

fighting their enemy in a long war.

正在一场持久的战争之中抗敌。

Now, it happened so one day:

有一天，发生了这样的故事:

David, I am not feeling very well today.

大卫，我今天身体不太舒服。

Would you please take these to your brothers?

能请你帮我把这些带给你的哥哥们吗?

Of course, father!

当然，父亲!

I will be glad to take food to my brothers.

我很乐意给我的哥哥们带去食物。

Please take rest!

请好好休息吧!

So, David set off from his house

所以，大卫从家里出发了，

and soon reached the place where the battle was taking place.

并很快到达了战场。

The sun had started to set,

太阳开始落山，

and as a rule, both the armies **retreat**ed.

作为规则，双方的军队都撤退了。

You see, they would only fight during the day

你瞧，他们只会在白天战斗，

and **retreat** when the sun would go down.

在太阳落山时就会撤退。

David! What are you doing here?

大卫！你在这里干什么?

I am here with your food.

我来这里送食物。

Why would you bring us food? Where is father?

为什么是你来给我们送食物？父亲呢？

He is here to make fun of our situation.

他是来嘲笑我们的处境的。

We are tirelessly fighting a war,

我们正在无休止地打这场仗，

while he relaxes all day in the fields, watching his sheep.

而他整天都在地里悠悠哉哉，放着他的羊。

Brother, actually –

哥哥们，事实上——

Enough! Give that to us and be gone from here!

够了！把东西给我们然后离开这里！

David felt sad that his brothers misunderstood him,

大卫为他的哥哥们误解了他而感到伤心，

but knew that they were tired and stressed because of the war.

但是他也知道哥哥们正因为这场战争而既疲惫又紧张。

He decided to leave.

于是，他决定离开。

But just then, the ground **trembl**ed and a shadow cast over him.

但就在那时，大地颤抖，一个阴影盖过了他。

David looked up to see a huge man, nine feet tall,

大卫抬起头看，只见一个九尺高的巨大的男人，

wearing a steel **armour** and a thick helmet.

穿着一身钢盔甲，戴着一顶厚厚的头盔。

He also carried huge weapons in his hands,

这个巨人手中还拿着巨大的武器，

and looked at their army with anger and gritted teeth.

正愤怒地，咬牙切齿地看着他们的军队。

Who, who is he?

他，他是谁？

While David’s brother would not entertain him,

虽然大卫的哥哥不想招待他，

a young soldier felt bad for him

一个年轻的士兵可怜他，

and started to tell him all about the huge man.

便开始告诉他关于这个巨人的事情。

His name is Goliath. He is a **bully**.

他的名字是歌利亚。他是一个恶霸。

Since the last forty days,

在过去的四十天里，

he has been showing up during the afternoon,

他每天下午都出现在战场上，

and challenging us for a duel.

向我们发起决斗的挑战。

Seeing his size, none of our soldiers have stepped forward.

看到他的体型，我们没有士兵敢走上前。

He is using this to his advantage, and so is their army.

他和他的军队都利用这一点占尽了优势。

Another day goes by,

又一天过去了，

and yet, no one from your army accepts my challenge.

然而，你们军队还是没有人接受我的挑战。

I will not wait longer.

我不会再等下去了。

My patience is wearing off.

我的耐心在一点点消失。

One of the days, I will pick one of you

我会在某天从你们之中挑出一个人，

and we shall fight.

来跟我决斗。

Huh!

喝！

Saying so, the giant went away.

巨人边说着，边走开了。

David looked at the soldiers of his country

大卫看着他的国家的士兵们，

and felt very sad.

感到十分难过。

No one should be this **intimidating**!

没有人可以如此令人生畏，盛气凌人！

What can we do?

我们该怎么办？

Trust God and take a leap of **faith**.

相信神明，他能让我们进行信仰之跃。

What do you mean?

你这是什么意思？

Emm…

嗯……

That evening, as the soldiers went back to their tents,

那天晚上，当士兵们都回到了帐篷时，

David went to speak with the King.

大卫去找国王谈话。

Your Highness! My name is David,

陛下！我是大卫，

and my brothers are soldiers in your army.

我的哥哥们都是您军队中的士兵。

OK? And?

哦？然后呢？

I have come here to ask you to **grant** me a wish.

我来这里是想请您满足我一个愿望。

Listen kid. Look around.

听着孩子。看看我们国家的样子。

I am not exactly in the situation to **grant** wishes.

现在不是实现你愿望的时候。

I ask for your permission to let me fight Goliath.

我是来向您申请与歌利亚决斗的许可的。

The King was shocked,

国王十分震惊，

his mouth opening wide.

惊讶到嘴巴都咧开了。

You? Are you out of your mind?

你？你疯了吗？

With God by my side,

有神明在我左右，

I fear no one.

我不惧一切。

The King thought for a while.

国王沉思了片刻。

He really wanted someone to fight Goliath.

他真的很想有人能和歌利亚决斗。

But David? He thought.

但是大卫？他想。

Well, alright!

啊，好吧！

But I am sure you are aware of the **inevitable** consequence.

但我肯定你清楚这么做会引起什么不可避免的结果。

David said nothing,

大卫没有说话，

but simply bowed down and left.

只是鞠了一躬，离开了。

And then came the next morning.

时间来到了第二天早上。

The soldiers tried to put **armour**s on David,

士兵们试着给大卫穿上盔甲，

but none fit him.

但没有适合他的。

It is fine. God is my **armour**.

这没关系。神就是我的盔甲。

And saying so, he left the tent

这么说着，他走出了帐篷，

and went to the nearby stream.

去了旁边的小溪。

There, he chose a few stones to use with his slingshot,

在小溪边，他选了几个石头用于他的弹弓，

and then looked up at the sky and prayed.

然后看向天空，祈祷着。

In your name, God! In your name!

以你的名义，神啊！以你的名义！

And then, he went to the battlefield,

然后，他来到了战场上，

where impatient Goliath awaited him.

而不耐烦的歌利亚已经在等他了。

What? What is this?

什么？这是什么？

I expected your strongest man

我以为来的会是你们最强的人，

and you send me a stick?

结果你们送来了一个竹竿般瘦弱的人？

Is this a joke?

你们在开玩笑吗?

Do not be so **arrogant**

别这么傲慢,

as to think of yourself as the most powerful.

把你自己当成是最强大的。

For only God is the most powerful of all.

因为只有神是万物中最强大的。

Let's see!

那我们走着瞧!

Goliath drew out the sword from his scabbard

歌利亚从剑鞘中抽出宝剑,

and roared with all his might

用他全部的力量大声咆哮,

trying to intimidate David.

想试图吓住大卫。

But David did not fear him.

但大卫并不害怕他。

Instead, he took out his stones,

与之相反,他拿出了石头,

put them in the slingshot,

把它们放到弹弓里,

and then swinging it,

甩着弹弓,

hurled it towards the giant.

用力地将石子投向了巨人。

Goliath could not even wink,

歌利亚甚至连眨眼的功夫都没有,

as the stones came hurtling towards him like the wind

石头如风一般猛地砸向他,

and hit him right in the forehead.

正中他的额头。

He fell on the ground, face first.

他脸部着地，摔到了地上。

Huh?

啊？

Unbelievable!

不敢相信！

Everyone was shocked,

所有人都大吃一惊，

and the enemy,

而敌人们，

seeing the unthinkable fold in front of them, fled.

看到了这一不可思议的信徒，逃跑了。

Everyone cheered and carried David on their shoulders.

所有人都在欢呼，他们把大卫扛到了肩膀上。

Thank you, David!

谢谢你，大卫！

You are indeed the bravest!

你确实是最勇敢的！

Do not thank me, Oh King, but the Lord,

哦，国王，不要感谢我，要感谢神，

for it is him who has defeated the enemy.

因为是他击败了敌军。

I was only but a medium.

我只是他施展力量的介质。

The King was moved by David's **devotion** towards God.

国王被大卫对神的虔诚感动了。

It isn't easy for a king to say that an average citizen of his country

这对一个国王来说绝非易事——让他说自己的一个普通子民

has taught him a great lesson.

给他好好上了一课。

But I am not one of them.

但我不是这样的国王。

You, young David, have taught me, and all of us that

你，年轻的大卫，给我和我们所有人上了一课，

belief gives us hope,

是你告诉了我们信仰给我们希望，

belief gives us strength,

信仰给我们力量，

and above all,

并且最重要的是，

belief gives us the courage

信仰给了我们勇气，

to face the toughest of the tough.

来面对最艰难的险阻。

Video B

What is Peking Opera?

京剧是什么？

Peking Opera is a classic form of Chinese performance art,

京剧是中国表演艺术中的一个传统形式，

which **combine**s the music of Chinese instruments,

它结合了中国乐器的音乐，

singing, miming, dancing, and history.

歌唱、哑剧、舞蹈和历史。

Peking Opera arose in the late 18th century,

京剧始源于 18 世纪末，

and by the 19th century,

到 19 世纪时，

the art form was popular throughout the country.

这个艺术形式已经风靡全国。

Today, it is regarded as

今天，它被认定为

one of China's cultural treasures.

中国的文化遗产之一。

There are four main characters in Peking Opera.

京剧中有四个主要角色。

This mask represents the character Sheng.

这个脸谱代表着角色“生”。

In addition, there are also three other major roles:

除此之外，还有另外三个主要角色：

Dan, Jing and Chou.

旦，净，和丑。

Each character has a different age,

每一个角色都有不同的年龄，

gender, and social status.

性别，和社会地位。

Sheng is the main male role in Peking Opera.

生是京剧中主要的男性角色。

Dan is a female character.

旦是女性角色。

However, interestingly,

然而，有趣的是，

young male performers actually played this role,

演这个角色的实际上是年轻的男演员，

as women were prohibited from acting until 1912.

而女性在 1912 年前都被禁止上台演出。

In Mandarin, ‘Jing’ means clean.

在普通话中，“净”的意思是干净。

The face paint does not make the character appear clean,

脸谱的绘画却并没有让角色看起来很干净，

so the word ‘Jing’ is used in a **contradictory** sense.

所以“净”这个词的用法是矛盾的。

Chou means ugly.

“丑”的意思是丑陋。

The actors who play the chou roles –

扮演丑角的演员——

although they may not necessarily be unattractive physically,

虽然他们并不一定长相丑陋，

have a white dot drawn on their noses.

但是他们会在鼻子上画上一个白点。

This feature means they are supposed to be ugly.

这一特征意味着他们应该是丑陋的。

The masks in Peking Opera represent

京剧中的脸谱代表了

the different characteristics of the personas on stage.

台上的角色们的不同特征。

As red is traditionally a positive colour in Chinese culture,

因为红色在中国传统文化中是一个代表正面意义的颜色，

The culture signifies **prosperity**, intelligence and bravery.

中国文化重视繁荣，智慧和勇气。

Purple masks suggest justice and **sophistication**.

紫色的脸谱代表着正义和复杂。

Although black is depicted as a negative colour in Western culture,

虽然黑色在西方文化中是一个负面的颜色，

the black mask is regarded as **neutral**,

但是黑色的脸谱被认为是中性的，

whereas the green masks and yellow masks

然而绿色和黄色的脸谱

represent bad characteristics.

代表着负面的特征。

The green mask implies violence,

绿色的脸谱暗示着暴力，

and the yellow cruelty.

而黄色代表了残忍。

Similarly, although Westerners traditionally

同样地，虽然西方人通常

see white as a symbol of purity,

将白色看作纯洁的标志，

the white mask actually stands for evil.

但是白色的脸谱实际上代表着邪恶。

Peking Opera actors have to train from early childhood

京剧的演员必须从很小的时候就开始训练，

as they need to master almost all the performing arts,

因为他们需要掌握几乎所有的表演艺术，

like singing, dancing and acting.

例如歌唱，舞蹈和表演。

Similarly, without perfect **martial** arts skills,

同样地，如果没有完美的武术技能，

it would be very difficult for an individual to become a Peking Opera professional.

一个人想要成为职业的京剧演员非常困难。

There are very strict rules actors must follow.

演员们需要遵守非常严格的规则。

During the 19th century,

在 19 世纪，

young pupils often lived in

年幼的小学生经常住在

special schools teaching Peking Opera.

教京剧的特殊学校里。

As the school provided all food and accommodation,

因为学校免费提供所有的食物和住宿，

the students were **indebt**ed to their trainers,

学生们受了师傅的恩惠，对他们有所亏欠，

whom they would repay once they made it as performers.

而他们一旦成为了演员，他们就会偿还和报答师傅。

Daily life was **exhausting** and brutal.

每一天都是精疲力竭和残酷的。

It was not uncommon for classes to run from 5 a.m. until the evening.

课程从早上五点上到晚上并非罕见。

Peking Opera is a virtual art,

京剧是一个虚拟的艺术，

and therefore not as straightforward as films and plays.

因此它不像电影和戏剧那样直观。

For example, the stage is meant to represent

比如说，舞台是展现了

a totally different reality.

一个完全不同的现实。

So if the actor walks in circles on the stage,

所以如果演员们在台上绕着圈走路，

it could symbolize that he or she

这可能代表着他/她

has walked miles and miles.

走了很长很长的路。

Similarly, the actors use a whip to hit an imaginary horse

同样，演员们用鞭子抽打一个想象中的马

to pretend that they are **gallop**ing across the land.

来假装他们在地上飞奔驰骋。

Basically, this is Peking Opera.

总的说来，这就是京剧。

Appendix E Vocabulary Levels Test

0-1000 Level

	choice (1)	computer (2)	garden (3)	photograph (4)	price (5)	week (6)
cost (1)	○	○	○	○	○	○
picture (2)	○	○	○	○	○	○
place where things grow outside (3)	○	○	○	○	○	○

	eye (1)	father (2)	night (3)	van (4)	voice (5)	year (6)
body part that sees (1)	○	○	○	○	○	○
parent who is a man (2)	○	○	○	○	○	○
part of the day with no sun (3)	○	○	○	○	○	○

	center (1)	note (2)	state (3)	tomorrow (4)	uncle (5)	winter (6)
brother of your mother or father (1)	○	○	○	○	○	○

middle (2)	○	○	○	○	○	○
short piece of writing (3)	○	○	○	○	○	○

	box (1)	brother (2)	horse (3)	hour (4)	house (5)	plan (6)
family member (1)	○	○	○	○	○	○
sixty minutes (2)	○	○	○	○	○	○
way of doing things (3)	○	○	○	○	○	○

	animal (1)	bath (2)	crime (3)	grass (4)	law (5)	shoulder (6)
green leaves that cover the ground (1)	○	○	○	○	○	○
place to wash (2)	○	○	○	○	○	○
top end of your arm (3)	○	○	○	○	○	○

	drink (1)	educate (2)	forget (3)	laugh (4)	prepare (5)	suit (6)

get ready (1)	○	○	○	○	○	○
make a happy sound (2)	○	○	○	○	○	○
not remember (3)	○	○	○	○	○	○

	check (1)	fight (2)	return (3)	tell (4)	work (5)	write (6)
do things to get money (1)	○	○	○	○	○	○
go back again (2)	○	○	○	○	○	○
make sure (3)	○	○	○	○	○	○

	bring (1)	can (2)	reply (3)	stare (4)	understand (5)	wish (6)
say or write an answer to somebody (1)	○	○	○	○	○	○
carry to another place (2)	○	○	○	○	○	○
look at for a long time (3)	○	○	○	○	○	○

	alone (1)	bad (2)	cold (3)	green (4)	loud (5)	main (6)
most important (1)	○	○	○	○	○	○

not good (2)	○	○	○	○	○	○
not hot (3)	○	○	○	○	○	○

	awful (1)	definite (2)	exciting (3)	general (4)	mad (5)	sweet (6)
certain (1)	○	○	○	○	○	○
usual (2)	○	○	○	○	○	○
very bad (3)	○	○	○	○	○	○

1000-2000 Level

	coach (1)	customer (2)	feature (3)	pie (4)	vehicle (5)	weed (6)
important part of something (1)	○	○	○	○	○	○
person who trains members of sports teams (2)	○	○	○	○	○	○
unwanted plant (3)	○	○	○	○	○	○

	average (1)	discipline (2)	knowledge (3)	pocket (4)	trap (5)	vegetable (6)
food grown in gardens (1)	○	○	○	○	○	○

information which a person has (2)	○	○	○	○	○	○
middle number (3)	○	○	○	○	○	○

	circle (1)	justice (2)	knife (3)	onion (4)	partner (5)	pension (6)
round shape (1)	○	○	○	○	○	○
something used to cut food (2)	○	○	○	○	○	○
using laws fairly (3)	○	○	○	○	○	○

	cable (1)	section (2)	sheet (3)	site (4)	staff (5)	tank (6)
part (1)	○	○	○	○	○	○
place (2)	○	○	○	○	○	○
something to cover a bed (3)	○	○	○	○	○	○

	apartment (1)	cap (2)	envelope (3)	lawyer (4)	speed (5)	union (6)
cover for letters (1)	○	○	○	○	○	○
kind of hat (2)	○	○	○	○	○	○
place to live inside	○	○	○	○	○	○

a tall building (3)						

	argue (1)	contribute (2)	quit (3)	seek (4)	vote (5)	wrap (6)
cover tightly and completely (1)	○	○	○	○	○	○
give to (2)	○	○	○	○	○	○
look for (3)	○	○	○	○	○	○

	avoid (1)	contain (2)	murder (3)	search (4)	switch (5)	trade (6)
have something inside (1)	○	○	○	○	○	○
look for (2)	○	○	○	○	○	○
try not to do (3)	○	○	○	○	○	○

	bump (1)	complicate (2)	include (3)	organize (4)	receive (5)	warn (6)
get something (1)	○	○	○	○	○	○
hit gently (2)	○	○	○	○	○	○
have as part of something (3)	○	○	○	○	○	○

	available (1)	constant (2)	electrical (3)	medical (4)	proud (5)	super (6)
feeling good about what you have done (1)	○	○	○	○	○	○
great (2)	○	○	○	○	○	○
happening all the time (3)	○	○	○	○	○	○

	environmental (1)	junior (2)	pure (3)	rotten (4)	smooth (5)	wise (6)
bad (1)	○	○	○	○	○	○
not rough (2)	○	○	○	○	○	○
younger in position (3)	○	○	○	○	○	○

2000-3000 Level

	angle (1)	apology (2)	behavior (3)	bible (4)	celebration (5)	portion (6)
actions (1)	○	○	○	○	○	○
happy occasion (2)	○	○	○	○	○	○
statement saying	○	○	○	○	○	○

you are sorry (3)						

	anxiety (1)	athlete (2)	counsel (3)	foundation (4)	phrase (5)	wealth (6)
combination of words (1)	○	○	○	○	○	○
guidance (2)	○	○	○	○	○	○
large amount of money (3)	○	○	○	○	○	○

	agriculture (1)	conference (2)	frequency (3)	liquid (4)	regime (5)	volunteer (6)
farming (1)	○	○	○	○	○	○
government (2)	○	○	○	○	○	○
person who helps without payment (3)	○	○	○	○	○	○

	asset (1)	heritage (2)	novel (3)	poverty (4)	prosecution (5)	suburb (6)
having little money (1)	○	○	○	○	○	○
history (2)	○	○	○	○	○	○
useful thing (3)	○	○	○	○	○	○

	audience (1)	crystal (2)	intelligence (3)	outcome (4)	pit (5)	welfare (6)
ability to learn (1)	○	○	○	○	○	○
deep place (2)	○	○	○	○	○	○
people who watch and listen (3)	○	○	○	○	○	○

	consent (1)	enforce (2)	exhibit (3)	retain (4)	specify (5)	target (6)
agree (1)	○	○	○	○	○	○
say clearly (2)	○	○	○	○	○	○
show in public (3)	○	○	○	○	○	○

	accomplish (1)	capture (2)	debate (3)	impose (4)	proceed (5)	prohibit (6)
catch (1)	○	○	○	○	○	○
go on (2)	○	○	○	○	○	○
talk about what is correct (3)	○	○	○	○	○	○

	absorb (1)	decline (2)	exceed (3)	link (4)	nod (5)	persist (6)
continue to happen (1)	○	○	○	○	○	○
goes beyond the limit (2)	○	○	○	○	○	○
take in (3)	○	○	○	○	○	○

	approximate (1)	frequent (2)	graphic (3)	pale (4)	prior (5)	vital (6)
almost exact (1)	○	○	○	○	○	○
earlier (2)	○	○	○	○	○	○
happening often (3)	○	○	○	○	○	○

	consistent (1)	enthusiastic (2)	former (3)	logical (4)	marginal (5)	mutual (6)
not changing (1)	○	○	○	○	○	○
occurring earlier in time (2)	○	○	○	○	○	○
shared (3)	○	○	○	○	○	○

Appendix F Comprehension Test

Video A

这个故事是关于什么的?

What is this story about?

战争 爱情 奇幻 家庭

大卫的职业是什么?

What is David's job?

牧羊人 士兵 国王 农民

大卫的哥哥们的职业是什么?

What are David's brothers' jobs?

牧羊人 士兵 国王 农民

大卫在上战场前带了什么?

What did David brought to the battle field?

盔甲 头盔 剑 弹弓

谁取得了战斗的胜利?

Who won the fight?

大卫 巨人 哥哥 国王

Video B

这部视频是关于什么的?

What is this video about?

京剧 电影 黄梅戏 二胡

京剧不包含哪个内容?

Which does Peking Opera NOT include?

乐器 唱歌 跳舞 太极

以下哪一个不是京剧中的主要角色

Which of these is NOT a main character in Peking Opera?

生 净 末 丑

丑的脸谱有什么特征?

What characteristic does Chou's face have?

鼻子上有白点 全脸涂白 全脸涂黑 耳朵涂黑

在舞台上转圈意味着什么?

What does it mean when the actor walk in circles on the stage?

走了很长的路 迷路 角色死亡 角色在追逐人

Appendix G Word Form Test

Choose the word you heard in the video.

Video A

1
faint
feeble
faith
fault

2
grant
grand
blunt
guilt

3
rewind
refrain
retreat
revolute

4
evident
inevitable
profitable
indifferent

5
devotion
devaluation
devastation
destroy

6
crumble
tremble
rumble
translate

7
archer
administer
armour
mourn

8
bully
fully
belly
bracket

9
imitating
intervening
frightening
intimidating

10
arrange
arrogant
arouse
intelligent

Video B

1
convince
combine
covariate
rewind

2
explanatory
haunting
exhausting
tiring

3
genetic
gender
tender

render

4
contradictory
predatory
construing
comprehending

5
presentation
provocation
serendipity
prosperity

6
comprehension
sophistication
complication
investigation

7
neural
neutral
mutual
nutritional

8
magistrate
martial
marginal
magical

9
index
indebt
indie
invest

10
garment
grasp
gown
gallop

Appendix H Word Meaning Test

Choose the correct Chinese meaning of the English word.

Video A

1
faith
决心
信念，信仰
虚弱的
失败

2
grant
授予
巨大的
头等的，最高的
收回

3
retreat
回答
请客
被击败
撤退

4
inevitable
可以食用的
必然的，不可避免的
神奇的，伟大的
恐怖的

5
devotion
忠诚
投票
进化
革命

6

tremble
寺庙
发抖
鼓起勇气
放弃

7
armour
盔甲
武器
头盔
子弹

8
bully
公牛
恶霸
战士
巨人

9
intimidating
吓人的
模仿的
胆小的
亲密的

10
arrogant
直率的
傲慢的
勇猛的
安排好的

Video B

1
combine
结合
分离
复杂化
引导

2

exhausting
轻松愉悦的
努力的
不爽的
筋疲力尽的

3
gender
年龄
职业
性别
婚姻状态

4
contradictory
一致的
矛盾的
复杂的
简单的

5
prosperity
破败
繁荣
成就
合理

6
sophistication
理解
理性
简单
复杂

7
neutral
复杂的
正面的
负面的
中立的

8
martial
边缘的

博弈的
武术的
对抗的

9
indebt
使……受恩惠，亏欠
使……失败
使……成功
使……联系在一起

10
gallop
散步
飞奔
竞走
静止

Appendix I Video Feedback Questions

请对这个视频与后面配套测试的难度进行反馈

	非常简单	很简单	简单	适中	难	很难	非常难
	1 (1)	2 (2)	3 (3)	4 (4)	5 (5)	6 (6)	7 (7)
视频内容的难度	○	○	○	○	○	○	○
阅读理解的难度	○	○	○	○	○	○	○
词汇测试的难度	○	○	○	○	○	○	○

你觉得视频的速度如何?

	特别快	很快	快	适中	慢	很慢	特别慢	
	1 (1)	2 (2)	3 (3)	4 (4)	5 (5)	6 (6)	7 (7)	
快	○	○	○	○	○	○	○	慢

你对于这个视频和相关测试的内容还有什么想说的吗（可不答）

_____

Appendix J CUREC Approval

Dear Yaguang Li

31 March 2022 17:26

Title: The effect of English videos with different types of subtitles and glosses on vocabulary learning of Chinese ESL learners

Ref: CIA-22HT-043

The above application has been considered on behalf of the Departmental Research Ethics Committee (DREC) in accordance with the procedures laid down by the University for ethics approval of all research involving human participants.

I am pleased to inform you that, on the basis of the information provided to DREC, the proposed research has been judged as meeting appropriate ethical standards, and accordingly, approval has been granted.

If your research involves participants whose ability to give free and informed consent is in question (this includes those under 18 and vulnerable adults), then it is advisable to read the following NSPCC professional reporting requirements for cases of suspected abuse
http://www.nspcc.org.uk/globalassets/documents/information-service/factsheet-child-abuse-reporting-requirements-professionals.pdf

Should there be any subsequent changes to the project which raise ethical issues not covered in the original application you should submit details to research.office@education.ox.ac.uk for consideration.

You probably know about this meta-analysis already, but in case you don't: Perez, M.M., Van Den Noortgate, W., Desmet, P. (2013). Captioned Video for L2 listening and vocabulary learning : A meta-analysis. System, 41 (3), 720-739.

Good luck with your research study.

Best wishes
Hamish Chalmers
Member of the DREC