Research Article

Andi Wang* and Ana Pellicer-Sánchez

Exploring L2 learners' processing of unknown words during subtitled viewing through self-reports

Abstract: Studies have shown the benefits of subtitled viewing for incidental vocabulary learning, but the effects of different subtitling types varied across studies. The effectiveness of different types of subtitled viewing could be related to how unknown vocabulary is processed during viewing. However, no studies have investigated L2 learners' processing of unknown words in viewing beyond exploring learners' attention allocation. The present research followed a qualitative approach to explore L2 learners' processing of unknown words during subtitled viewing under three conditions (i.e., captions, L1 subtitles, and bilingual subtitles) by tapping into learners' reported awareness of the unknown words and the vocabulary processing strategies used to engage with unknown words. According to stimulated recall data (elicited by eye-tracking data) from 45 intermediate-to-advanced-level Chinese learners of English, captions led to increased awareness of the unknown words. Moreover, the types of strategies learners used to cope with unknown vocabulary were determined by subtitling type.

*Corresponding author: Andi Wang, National Research Centre for Foreign Language Teaching Materials, School of English and International Studies, Beijing Foreign Studies University, No. 2 Xisanhuan North Road, Haidian District, Beijing, China, E-mail: andi.wang@bfsu.edu.cn. https://orcid.org/0000-0002-7228-1150

Ana Pellicer-Sánchez, IOE, UCL's Faculty of Education and Society, University College London, London, UK, E-mail: a.pellicer-sanchez@ucl.ac.uk. https://orcid.org/0000-0002-9401-7027

1 Introduction

Numerous studies have documented the benefits of watching foreign language audio-visual materials (e.g., watching films and television series; henceforth viewing) for second language (L2) vocabulary learning (e.g., Peters and Webb 2018). The advantages of viewing for learning have been supported by Paivio's (1986) dual-coding theory and the multimedia principle in Mayer's (2009) cognitive theory of multimedia learning. They proposed that learning outcomes can be enhanced when learners' separate verbal and visual channels are both activated. In viewing, the dynamic images are processed via the visual channel while the audio soundtrack is processed through the verbal channel. This integration of visual and verbal information is believed to facilitate viewers' understanding, leading to greater depth of processing, and better information retention (Montero Perez et al. 2013).

The use of on-screen text such as captions (i.e., transcripts written in the same language as the text spoken in the video), first language (L1) subtitles (i.e., transcripts translated into the viewer's L1), and bilingual subtitles (i.e., simultaneous presentation of L1 subtitles and captions) has been found to further enhance this benefit (Montero Perez 2022; Wang and Pellicer-Sánchez 2022). Studies comparing the effectiveness of different subtitling types have reported the superiority of captions over other subtitling types for facilitating the learning of word form (e.g., Peters 2019), whereas bilingual subtitles seem to be more effective for facilitating knowledge of word meaning (e.g., Li and Hennebry-Leung 2022; Wang and Pellicer-Sánchez 2022). The varied learning gains measured by different vocabulary tests seem to suggest that the various subtitling types make learners engage with unknown vocabulary differently. However, most research has focused on examining learning gains without exploring learners' use and processing of unknown vocabulary in different types of on-screen text.

A few recent eye-tracking studies have shown that the amount of attention paid to L2 unknown words varied across subtitling types, and more attention to the L2 unknown words did not always lead to higher learning gains (e.g., Montero Perez et al. 2015; Wang and Pellicer-Sánchez 2022). These studies, while providing useful information about the amount of attention paid to unknown vocabulary during subtitled viewing, were not able to shed light into the various underlying cognitive processes that learners engage with during subtitled viewing, as the examination of eye movements does not allow for the differentiation between learning effort, processing difficulty, or superficial viewing behaviour (Wang and Pellicer-Sánchez 2022). Researchers have thus called for the use of qualitative data to further investigate learners' cognitive processes in subtitled viewing (Montero Perez et al. 2015).

One important aspect that has been studied in the processing of unknown vocabulary is learners' awareness of the unknown words (Godfroid and Schmidtke 2013). Previous research has shown that advanced L2 learners were less aware of the pseudowords during L2 reading, and awareness was a strong predictor of word recognition. However, this research has been conducted in the context of learning from reading, and similar investigations into learning from viewing are yet to be conducted. Another important aspect of the processing of unknown vocabulary relates to the vocabulary processing strategies employed (e.g., Fraser 1999; Rott 2005). To date, little research has investigated L2 learners' processing strategies in vocabulary learning in general, especially in the context of learning from viewing. To the best of our knowledge, Sydorenko (2010) is the only empirical study that has explored learners' vocabulary learning strategies during viewing. However, this research only investigated the general strategies used during captioned viewing. It is possible that the findings did not represent what happened with each individual word. Therefore, our understanding of the processing strategies employed during the various types of subtitled viewing is still limited.

The present study addresses these gaps by investigating L2 learners' processing of unknown vocabulary during different subtitled viewing conditions (captions, L1 subtitles, bilingual subtitles), focusing on learners' reported awareness of unknown words and vocabulary processing strategies. Importantly, this study addresses the lack of qualitative studies in vocabulary learning research (Webb 2020) and responds to the call for qualitative investigations to examine cognitive processes during viewing (Montero Perez et al. 2015).

2 Literature review

2.1 Incidental vocabulary learning from subtitled viewing

Viewing has been advocated as a valuable type of out-of-class language input that can increase L2 learners' exposure to authentic L2 input (Vanderplank 2010), which also facilitates their incidental vocabulary learning (Peters and Webb 2018). In the current research, incidental vocabulary learning is defined from a methodological perspective as not informing participants of the existence of vocabulary tests before the experiment (e.g., Peters and Webb 2018; Peters et al. 2016). The use of on-screen text seems to be more effective for incidental vocabulary learning compared to viewing without on-screen text (e.g., Koolstra and Beentjes 1999; Li 2016; Wang and Pellicer-Sánchez 2022). However, research comparing the effects of different subtitling types on incidental vocabulary learning has yielded inconclusive findings. Most studies have used pre- and post-vocabulary tests to assess learners'

knowledge of the target words in the video before and after viewing, and these tests mainly targeted learners' knowledge of the word form and/or meaning. These studies, focusing on the examination of learners' performance on vocabulary tests, have shown a tendency for captions to be more effective than L1 and bilingual subtitles for learning word form (e.g., Peters 2019; Peters et al. 2016; Wang and Pellicer-Sánchez 2022), but their effectiveness seems to be similar to L1 subtitles for learning word meaning (e.g., Pujadas and Muñoz 2019; Peters et al. 2016). The effect of bilingual subtitles is still controversial, as some studies showed their advantages over captions and L1 subtitles for facilitating meaning knowledge (e.g., Li and Hennebry-Leung 2022; Wang and Pellicer-Sánchez 2022), while others revealed no significant differences (e.g., Lwo and Lin 2012). Researchers have attempted to explain these findings from post-viewing vocabulary tests, speculating how the different subtitling conditions support incidental vocabulary learning. However, to date, no studies have provided empirical evidence to support these speculations by looking in-depth into how L2 learners actually use different on-screen texts to facilitate vocabulary learning.

A few recent studies have used eye-tracking to investigate learners' attention to unknown vocabulary during incidental learning from subtitled viewing. These studies have reported that textually enhanced words in captions received greater visual attention than unenhanced ones (Puimège et al. 2023), but that similar amount of attention was paid to unknown vocabulary in keyword and full captions in incidental learning conditions (Montero Perez et al. 2015). Inconclusive findings have been reported concerning the relationship between amount of attention and learning gains. The study by Puimège et al. (2023) suggested that more visual attention to unknown vocabulary seemed to lead to higher learning gains, whereas the study by Montero Perez et al. (2015) showed that longer first-pass reading time on unknown words led to higher form recognition gains, but longer second-pass reading time resulted in lower gains. The authors pointed out that longer reanalysis might not reflect increased intention to learn a word but rather processing problems. An eye-tracking study by Wang and Pellicer-Sánchez (2022) compared the effectiveness of bilingual subtitles for incidental vocabulary learning with that of captions, L1 subtitles, and no subtitles. The eye-tracking data suggested that learners using bilingual subtitles were less likely to process the L2 unknown words than those using captions, whereas they were more likely to process the corresponding L1 translations than those using L1 subtitles. Longer first-pass reading time and total reading time on the L2 word forms to some extent predicted the learning of word form and meaning, but this relationship was non-significant in second-pass reading time, suggesting an indirect relationship between the amount of attention paid to a word and the learning outcome.

These eye-tracking studies provided direct evidence for learners' visual attention to unknown words during subtitled viewing. Although eye-tracking is considered a reliable measure of learners' attention and unconscious processes, it cannot reveal learners' more conscious underlying cognitive processes (Godfroid 2020). Exploring these underlying processes is crucial to better understand the benefits of various subtitling types, which would in turn help researchers and practitioners maximise the benefits of subtitled viewing for vocabulary learning.

2.2 Qualitative approaches to explore the processing of unknown words in text and/or audio L2 input

Given the paucity of qualitative research probing into learners' processing of unknown words during viewing, a review of previous studies about learners' word processing in reading and listening conditions may inform us of the processes in viewing with on-screen text. While there have been some studies examining vocabulary learning strategies in intentional learning contexts (for a comprehensive review, see for example, Gu 2020; Schmitt 1997), in this section we focus on the few studies that have examined strategy use in vocabulary learning from natural exposure to written and auditory input. These studies have investigated L2 learners' cognitive processing of unknown words in L2 reading and listening by examining learners' reported awareness of the unknown lexical items and the processing strategies they employed when engaging with those items.

Godfroid and Schmidtke (2013) investigated L2 learners' processing of pseudowords and its relationship with learners' incidental learning gains by looking at both learners' attention (examined with eye-movement data) and awareness (examined with stimulated recalls). After the reading activity and the posttests, participants were asked to report their awareness of the pseudowords. Awareness in this study was coded into three categories: no awareness (participants did not consciously remember the word), noetic awareness (participants remembered the word was somewhere in the text), and autonoetic awareness (participants remembered the word in a particular sentence). The findings showed that participants were not aware of most of the pseudowords (67.4 %), and only reported some of the pseudowords with noetic (18.8 %) or autonoetic awareness (12.3 %). Both attention and awareness positively predicted word recognition, with awareness being the strongest predictor.

A few reading studies have investigated learners' vocabulary processing strategies used to construct their knowledge of unknown words during L2 reading (e.g., Rott 2000, 2005). Most of these studies used think-aloud protocols to explore L2 learners' strategy use by asking learners to verbalise their thoughts while reading, with or without meaning support (e.g., using L1 or a dictionary). Despite the different

coding used across L2 reading studies, some common strategies have emerged, including using context, morphological analysis, repeating aloud, skipping, using background knowledge, analogy, and self-inquiry (e.g., Fraser 1999; Hu and Nassaji 2012; Lawson and Hogben 1996; Nassaji 2003). Moreover, learners seem to use both a single strategy and a combination of various strategies when processing unknown words (e.g., Fraser 1999; Hu and Nassaji 2012; Rott 2000). These studies found that the most frequently used vocabulary processing strategies were inferring word meanings from context (e.g., Fraser 1999; Hu and Nassaji 2012; Lawson and Hogben 1996), and repeating out loud parts of the text including the target word (e.g., Lawson and Hogben 1996; Nassaji 2003). When meaning support was provided, L2 learners were also found to frequently use the L1 translations or dictionary to obtain word meanings (e.g., Fraser 1999; Lawson and Hogben 1996).

It should be noted that most of the reading studies above were set in an intentional vocabulary learning context where participants were deliberately asked or trained to infer the meaning of target words during reading. Therefore, findings might be different in incidental learning settings. Previous studies have indeed shown that learners tend to ignore unknown words in incidental learning conditions (Godfroid and Schmidtke 2013). Rott (2000) also found that in an incidental learning condition, skipping unknown words in L2 reading was a strategy used by both successful and less-successful L2 learners. However, a study by Prichard and Atkins (2021) revealed that when L1 definitions of novel words could be accessed during reading, participants were less likely to ignore the unknown words, and using L1 translations was the most frequently used strategy. These findings together suggest that in incidental learning settings, learners tend to ignore many of the unknown words, but when available, learners tend to use L1 translations as a shortcut for comprehension.

Most studies on processing unknown words have been conducted in the reading context, with very few in the L2 listening context. Vandergrift (2003) investigated the strategies that English learners of French used in listening comprehension, and found that learners used four types of strategies to guess the meaning of unknown words, including linguistic (i.e., known words in an utterance), voice (e.g., tone and/or paralinguistics), extralinguistic (i.e., background sounds and relationships between speakers in an oral text, material in the response sheet, or concrete situational referents), and between-parts (i.e., information beyond the local sentential level), with linguistic inferencing being more frequently used.

Cai and Lee (2012) explored Chinese EFL learners' awareness and processing strategies of nine pseudowords while listening to nine short audio recordings. The stimulated recall findings revealed that participants were not aware of about 24.4 % of the pseudowords. For those pseudowords that had been noticed, participants were

more likely to infer pseudoword meanings than to ignore them. When inferring word meaning, participants used local context in the pseudoword sentence or other sentences, overall context, background knowledge, paralinguistics (e.g., rhythm, intonation), phonology, word class, and morphology. Similar to reading, participants were also found to use a single strategy or a combination of strategies. Their most frequently used strategies were local context combined with background knowledge and overall context. However, no significant relationship between the vocabulary processing strategy used and word learning was observed. These findings suggested that the word processing strategies used in L2 listening were similar to those in reading, but learners may also use voice and paralinguistics to infer word meaning when audio input is available.

Sydorenko (2010) conducted the only empirical research to explore learners' vocabulary learning strategies during viewing. Beginner L2 learners of Russian were asked to watch three two-minute Russian videos twice (the first time focusing on comprehension and the second time on vocabulary learning) in one of three viewing conditions: video + audio + captions, video + audio, and video + captions. Participants' vocabulary learning strategies were examined using open-ended questions in a final questionnaire. Two types of general strategies emerged: modality-specific strategies (including matching visual images with words and reading captions) and common vocabulary guessing strategies (including recognising words that are similar to L1, using the roots of known words, paying attention to the verbal context, and paying attention to grammar). Using visual images was the most frequently used strategy in all viewing conditions. Participants in the video + audio + captions condition reported fewer use of common guessing strategies than participants in other two conditions. Two possible interpretations were provided: 1) little time was left for making inferences after processing all three input sources; 2) the use of three input sources was sufficient for participants to understand meanings of unknown words. The findings implied that strategies employed in captioned viewing seem to be different from those employed in reading and listening. However, participants were asked to focus on vocabulary learning during their second viewing. Thus, it is still unclear how learners engage with unknown words in a more natural viewing setting. Importantly, this research only investigated the overall strategies used during viewing using a questionnaire, which might not have captured all different strategies used for each individual word. In addition, no previous studies have investigated learners' vocabulary processing strategies in viewing with L1 or bilingual subtitles. It is not clear whether, when learners' L1 is available in the subtitling area, L2 learners would actively engage with L1 translations as reported in reading studies (e.g., Lawson and Hogben 1996; Prichard and Atkins 2021).

2.3 The present study

The review of the literature has shown that most studies on learning from subtitled viewing used offline vocabulary tests to compare the effects of different subtitling types on incidental vocabulary learning. While eye-tracking studies have provided a better understanding of the attention allocation to unknown vocabulary in various subtitling conditions, they have also revealed the need to further explore the underlying cognitive processes to provide a comprehensive understanding of the vocabulary learning process. L2 reading research has attempted to explore learners' processing of unknown vocabulary by examining learners' reported awareness and vocabulary processing strategies using qualitative methods. However, no research has thoroughly investigated this issue in subtitled viewing. The present study aimed to fill these gaps by investigating learners' awareness and processing strategies of unknown words during various subtitled viewing conditions.

Stimulated recall collects qualitative data by eliciting the thought processes that occur while a learner is doing a task/activity by asking the learner to verbalise those processes after the events with a prompt to stimulate their memory (Gass and Mackey 2017). Similar to think-aloud protocols, it gains information about learners' cognitive processes that take place during an event, with the particular advantage of being unobtrusive (Gass and Mackey 2017). We decided to use stimulated recalls to investigate participants' processing of the unknown vocabulary because viewing, where new information is continuously presented, does not allow participants to pause and verbalise their thoughts. It is believed that, if stimulated recall is conducted shortly after the task, thoughts are still in short-term memory and more valid information can be attained by cuing participants with specific material used in the experiment (Ericsson and Simon 1993; Gass and Mackey 2017).

The present research aimed to answer the following research questions (RQs):

1. To what extent are L2 learners aware of the unknown vocabulary during subtitled viewing?
2. Does subtitling type (i.e., captions, L1 subtitles, bilingual subtitles) lead to differences in learners' reported awareness of unknown vocabulary?
3. What strategies do L2 learners use to process the unknown vocabulary during subtitled viewing?
4. Does subtitling type (i.e., captions, L1 subtitles, bilingual subtitles) lead to differences in learners' use of vocabulary processing strategies?

3 Methods

The data examined in the present study are from a larger project on the effects of bilingual subtitles on various outcome measures. Results of offline tests and eye-tracking data are reported in Wang and Pellicer-Sánchez (2022, 2023). The present study focuses on the stimulated recall data. Thus, the participants and viewing materials are the same as those in the studies mentioned above.

3.1 Participants

Among the 112 participants who participated in the original study, data from 45 participants (15 per condition) were randomly selected for the present study. This sample size is considered sufficient, as a review of 88 stimulated recall studies found most research included fewer than ten participants, with only a few studies having between 30 and 77 participants (Sanchez and Grimshaw 2020). Participants were L1 Chinese speakers who were studying at a British university with various academic backgrounds. Their ages ranged between 18 and 26 years (M = 22.56 years, SD = 1.52, 95 % CI [22.10, 23.01]). Their average International English Language Testing System (IELTS) score was 6.73 (SD = 0.60, 95 % CI [6.55, 6.91]), which approximately corresponds to B2 to C1 levels in the Common European Framework of Reference for Languages according to the IELTS official guidelines (IELTS, n.d.). According to the background questionnaires, most of the participants used on-screen text while viewing in daily life, with bilingual subtitles ranked as the most frequently used subtitling type (M = 4.38, SD = 1.45, 95 % CI [3.94, 4.81], Max = 6), followed by captions (M = 3.13, SD = 1.29, 95 % CI [2.75, 3.52]), L1 subtitles (M = 3.04, SD = 1.30, 95 % CI [2.66, 3.43]), and no subtitles (M = 2.04, SD = 1.17, 95 % CI [1.69, 2.40]).

The 3,000-word-level Vocabulary Levels Test (Schmitt et al. 2001) was administered to ensure the comprehensibility of the selected viewing material. All the participants demonstrated their mastery of the 3,000-word level, suggesting their familiarity with most of the vocabulary in the selected material. All 45 participants could adequately understand the viewing material, with an average score of 78.88 % (SD = 13.59 %, 95 % CI [74.79 %, 82.97 %]) on the multiple-choice comprehension test (Cronbach's alpha at 0.83).

3.2 Materials

3.2.1 Viewing material

Four authentic video excerpts (in total 23 min, 3,488 words) from the BBC documentary *Animal Odd Couples* (BBC 2013) were chosen as the viewing material. They

were extracted and put together using the video editing software (Corel 2018). This documentary consists of several journeys taken by a wildlife biologist Liz Bonnin to find out why animals of different species develop unusual close relationships with each other. We analysed the video scripts using the Range software (Nation and Heatley 2002), with the British National Corpus (BNC Consortium 2007) as the reference corpus. The first 3,000 most frequent words provided a coverage of 96 % of the selected clips. The original English video script was retrieved online, and the first author translated the script into Chinese. To ensure the accuracy of the translation, it was compared to an online amateur translation (Bilibili n.d.) and checked by three L1 Chinese speakers fluent in English. The material was then piloted twice with 13 advanced Chinese learners of English. The on-screen text was added to the video using SrtEdit (Corel 2018; PortableSoft 2012) software and formatted according to the BBC *Subtitle Guidelines* (BBC 2019). In the bilingual subtitles condition, L1 and L2 lines were presented simultaneously with the L1 above L2 lines, in accordance with the common presentation format of bilingual subtitles in China. Three versions of the video were created, one for each subtitling condition (i.e., captions, L1 subtitles, bilingual subtitles). Example screenshots of the three subtitling conditions are presented in Appendix S1.

3.2.2 Target words

The original content and audio soundtrack of the video were used to maintain the ecological validity of the study. The following steps were taken to choose target words (henceforth TWs) that were potentially unknown for all participants: 1) we inspected the script and created a list of 66 potentially unknown words; 2) seven experienced Chinese IELTS teachers were asked to select the TWs from the list that they thought would be unknown to the participants; 3) 13 Chinese learners of English with similar characteristics to the participants were asked to indicate their knowledge of the remaining TWs. These resulted in a final selection of 24 single TWs. Since multiple exposures of a word could affect learners' engagement with the word (Schmitt 2008), and most TWs in the present study appeared only once in the video, we decided to focus on learners' initial processing of each TW and discarded the TWs that appeared more than once, resulting in a total of 19 TWs (seven nouns, eight verbs, and four adjectives; see Appendix S2 for details) included in the coding and analysis. The TWs that were known, as indicated in their vocabulary pretest or stimulated recalls, were further deleted from the analysis.

3.3 Procedure

After receiving the instructions and giving their signed consent, participants' prior knowledge of the TWs and their vocabulary size were examined via a series of tests (see Wang and Pellicer-Sánchez 2022 for details; see Appendix S3 for complete tests, also available at: https://www.iris-database.org) 2 to 3 weeks before the viewing. The 45 participants were randomly assigned to one of three subtitled groups (i.e., captions, L1 subtitles, and bilingual subtitles). They were asked to watch the viewing material on a 19-inch monitor with a 1920×1080 screen resolution individually in an eye-tracking lab. Participants were told to watch for comprehension, and they were not aware of the vocabulary posttests or stimulated recalls before the viewing. Their eye movements were recorded during viewing with EyeLink 1000 Plus (SR Research 2016), in desk-mounted mode. The participants were asked to wear headphones during the viewing session. After the viewing, all 45 participants were asked to complete a comprehension test and vocabulary posttests (see Appendix S4 for the results of these tests for the 45 participants in the present study, and Wang and Pellicer-Sánchez 2022, 2023 for the results corresponding to the whole pool of participants, $N = 112$).

Participants were not given the correct answers to the vocabulary posttests before conducting the stimulated recall interviews, which were individually held immediately after the vocabulary posttests. Oral instructions, adapted from Gass and Mackey (2017), were first given to each participant (see Appendix S5). Participants were encouraged to ask questions about the procedures. After instructions, participants' recall was prompted by the replay of their eye-movement recordings during the presentations of each TW (19 in total). Each participant was first asked if they were aware of each TW when they encountered them during subtitled viewing (i.e., "Did you notice this word at that time?"), and if they were, what they thought about at that time (i.e., "What were you thinking at that time when encountering this word during viewing?"). In the stimulated recall, *notice* was used as an everyday language, indicating learners' self-reported awareness of either the written or auditory form of TWs.

The 19 recordings were played at a 50 % speed using the EyeLink Data Viewer software package (SR Research, Version 3.1.246 2018). This decision was made after a pilot conducted with seven Chinese learners of English reported the difficulty in following their rapid eye movements to recall their thoughts. All stimulated recalls were held in participants' L1 (i.e., Mandarin Chinese) and were audio recorded with a portable recorder. The approximate average duration of each stimulated recall was 15 minutes. At the end of the experiment, participants were debriefed on the real purpose of this study and were given a small compensation for their participation.

3.4 Data analysis

The stimulated recall data from the 45 participants was first fully transcribed. Initially, 855 cases (45 participants × 19 TWs) were prepared for coding using NVivo 12 Pro software (QSR International Pty Ltd 2018). Since the focus was participants' processing of unknown words, those TWs that were familiar or partially familiar to each participant before the experiment, as indicated in the form recognition pretest results and their self-reports in the stimulated recall, were discarded from the analysis (268 cases in total; captions = 81 cases; L1 subtitles = 93 cases; bilingual subtitles = 94 cases).[1] This resulted in a final number of 587 cases of unknown words (captions = 204 cases; L1 subtitles = 192 cases; bilingual subtitles = 191 cases) for further coding.

All data were coded inductively following the steps suggested by Bryman (2012) and Selvi (2020): 1) concepts were generated by coding data at the level of open coding; 2) categories were generated through a constant comparison of concepts, micro-categories were grouped into more general categories; 3) saturated categories were listed; 4) categories were applied back to the stimulated recall data pertaining to each TW. As shown in Figure 1, to answer RQ1 and RQ3, the stimulated recall data was analysed at two levels. For RQ1, Level 1 (i.e., awareness level) was coded inductively to explore participants' reported awareness of each TW. Based on participants' answers to the first stimulated recall question (i.e., "Did you notice this word at that time?"), three categories emerged from the data: "forgot", "reported no awareness", and "reported awareness" (see Table 1). The coding did not distinguish
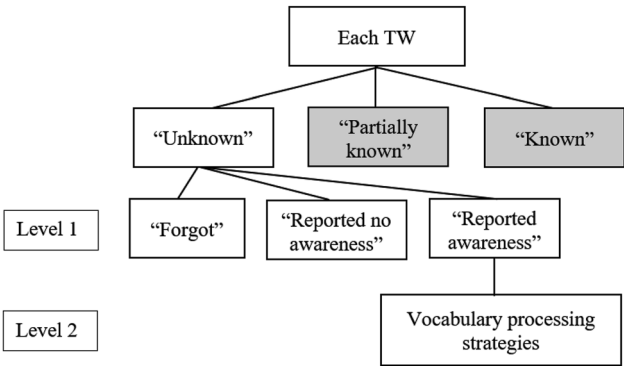

Figure 1: Illustration of stimulated recall coding procedures at two levels for each target word.

---

1 Form recognition pretest results were used to take into account partial knowledge and ensure that all words included in the analysis were fully unknown to participants.

Table 1: Categories of participants' reported awareness (Level 1) of unknown target words in stimulated recall data.

Categories	Definitions	Examples
Forgot	Forgot whether they were aware of the TW or not during viewing.	"I don't remember."
Reported no awareness	Reported no awareness of the TW, by explicitly mentioning their lack of awareness or by only commenting on video content, images, audio etc.	"I didn't notice this word at that time."; "I was working hard to understand the speakers' posture."
Reported awareness	Reported awareness of the TW; reported that they had seen/heard it or guessed its meaning during viewing.	"I noticed this word at that time."

whether participants were aware of the TWs in the auditory or written forms, due to the difficulty participants had in making this distinction when both forms were available. Definitions and examples of the coding categories at this level are presented in Table 1.

In response to the RQ2, to examine whether there were group differences concerning participants' reported awareness of the unknown TWs, the frequency and percentage for each code were calculated for each subtitling group. Since the homogeneity of variances assumption was violated ($p = 0.01$), a Welch ANOVA was run in SPSS (version 25) as an alternative to one-way ANOVA to explore the potential group differences, and Games-Howell test was used as an alternative post-hoc comparison.

To answer RQ3, only the TWs that were coded as "reported awareness" at Level 1 were coded for Level 2 (i.e., vocabulary processing strategies), as illustrated in Figure 1. To categorise the strategies used for those unknown TWs, the data were coded inductively following the steps suggested by Bryman (2012) and Selvi (2020) mentioned above. To name the categories emerging in the stimulated recall, previous work exploring L2 learners' vocabulary learning strategies (e.g., Lawson and Hogben 1996; Nation 2001) and vocabulary processing strategies in reading (e.g., Fraser 1999; Rott 2000, 2005) were used as references. It is important to note that the reported strategies were coded according to their content, regardless of whether they were incorrectly or inaccurately applied. To examine the reliability of the coding, 20 % of the data, i.e., data from nine randomly selected participants (three per group), were coded by an experienced second coder who was a Chinese speaker fluent in English. An interrater agreement of $\kappa = 0.96$ ($p < 0.001$, 95 % CI [0.93, 1]) at Level 1, and $\kappa = 0.89$ ($p < 0.001$, 95 % CI [0.82, 0.95]) at Level 2 was obtained. We then resolved

disagreements through subsequent discussion to reach 100 % interrater agreement on all items that were double coded.

To answer RQ4, we first calculated each participant's frequency of using each specific strategy. Since the number of noticed unknown words varied across subtitling conditions, we used the percentages of frequency (i.e., dividing each participant's frequency of using each strategy by the overall cases of strategy use in that participant's subtitling condition) to indicate how frequently each strategy was reported in a particular subtitling condition. Since the normality assumption was violated (all ps < 0.05), and given the small sample size, Kruskal–Wallis Tests, as a non-parametric alternative to one-way ANOVA, were run in SPSS to explore the potential group differences in participants' use of general vocabulary processing strategies. Post hoc comparisons were conducted using Mann–Whitney Tests with a Bonferroni adjusted alpha level of 0.017 (= 0.05/3).

4 Results

4.1 RQ1 and RQ2 – Level 1 awareness

RQ1 and RQ2 aimed to explore participants' awareness of the unknown words during subtitled viewing, and how that differed across subtitling conditions. As can be seen in Table 2, in general, there were slightly more cases of participants not being aware of the unknown TWs (53.49 %) than aware (40.72 %). However, the numbers varied across subtitling conditions. The captions group recorded the highest cases of reported awareness, followed by bilingual subtitles, and L1 subtitles, while the opposite was true for the cases of no awareness. It can also be observed that only the captions group reported more cases of awareness than no awareness.

The group difference was further confirmed by the Welch ANOVA analysis. There was a statistically significant difference in the percentage of the cases of

Table 2: Results for the frequency and percentage of participants' reported awareness (Level 1) of unknown target words in stimulated recalls by group.

Categories	Captions (%)	L1 (%)	Bilingual (%)	Total frequency (%)
Forgot	17 (8.33)	10 (5.21)	7 (3.66)	34 (5.79)
Reported no awareness	83 (40.69)	127 (66.15)	104 (54.45)	314 (53.49)
Reported awareness	104 (50.98)	55 (28.65)	80 (41.88)	239 (40.72)
Total (%)	204 (100)	192 (100)	191 (100)	587 (100)

reported awareness for the three groups, $F(2, 42) = 5.72$, $p = 0.03$, with a large effect size $\eta^2 = 0.21$ (Cohen 1988). Post-hoc comparisons indicated that the captions group reported significantly more cases of awareness of unknown TWs than L1 subtitles ($M = 0.24$, SE = 0.84, $p = 0.02$, 95 % CI [0.04, 0.45]). However, no significant differences emerged between bilingual subtitles and captions ($M = -0.12$, SE = 0.07, $p = 0.22$, 95 % CI [−0.29, 0.05]) or L1 subtitles ($M = 0.13$, SE = 0.06, $p = 0.14$, 95 % CI [−0.03, 0.29]).

4.2 RQ3 – Level 2 vocabulary processing strategies

RQ3 aimed to investigate the vocabulary processing strategies participants used to cope with the unknown TWs that they were aware of during subtitled viewing. As shown in Figure 1, only the TWs that were unknown and that participants had reported awareness were considered for the analysis, resulting in a total of 239 cases (captions = 104; L1 subtitles = 55; bilingual subtitles = 80). As Table 3 shows, there were six broad categories and 19 subcategories reported by the participants. Four broad categories were present in all subtitling conditions and included: 1. Word feature analysis, in which participants reported their processing of certain characteristics of the TW; 2. Using context, where participants used auditory, visual, and contextual information to support their understanding of the TW; 5. Others, including a few cases where participants mentioned specific strategies that did not fit in other broad strategies; and 6. No reported strategies, including those cases in which participants were aware of the TW but did not report any thoughts about the TW. The remaining two broad categories only applied to certain subtitling conditions: 3. Using L1 translations, where participants mentioned their use of the L1 translations, which was not applicable to the captions group; and 4. Guessing without reported sources, which did not apply to the L1 or bilingual subtitles groups because the presence of L1 translations made meaning guessing less likely to occur during viewing. The specific strategies and definitions are presented in Table 3.

It should be noted that in some cases, although participants first reported their awareness of an unknown TW, they did not report their thoughts about the TW. This was named as "6. No reported strategies". This category should be distinguished from the "Reported no awareness" category at Level 1 (awareness level), where participants did not report awareness of the TW.

In most cases, participants used only one strategy (89.12 %) for each unknown TW. However, in 26 cases (10.88 %), participants combined two (9.21 %) or three (1.67 %) strategies to process one TW. The number of strategies reported for the TWs that appeared at different times in the video did not differ greatly, with the average number of strategies for the TWs in the four parts of the video ranging from 1.11 to 1.14. Overall, there were 269 instances of strategy use for the 239 cases of unknown

Table 3: Categories of participants' vocabulary processing strategies (Level 2) for unknown target words in stimulated recalls.

Strategy categories	Subcategories	Further subcategories	Definitions	Example quote
1. Word feature analysis	1.1. Analysing part of speech	-	Analysed the part of speech or reported awareness of the part of speech of the TW	"Hmm I guessed it [*endearing*] was a positive adjective…"
	1.2. Analysing word-structure	-	Analysed the affixes, suffixes, and/or roots of the TW	"'*Confiscated*'… Because looking at the word structure, 'con' is very common, 'fiscated', I knew how to pronounce this type of word by a glance, but just didn't know its meaning."
	1.3. Form association	-	Associated the TW with a known word sharing similar form or pronunciation of the TW	"'*buffering*', I immediately thought about 'suffering'…"
	1.4. Analysing word pronunciation	-	Used the pronunciation of the TW to understand its meaning	"I thought it [*nuzzle*] means a nap, because the pronunciation of '*nuzzle*' sounded very intimate."
	1.5 Word usage	-	Reported the awareness of the usage of the TW	"'*buffering of stress*'. I heard this phrase. I had an impression of it at that time. I didn't know it could be used together with 'stress', so I thought about it."
2. Using context	2.1. Using auditory cues	-	Used the audio/sound of the video to support understanding of the TW	"Oh, I have a deep impression about this one, because the cat was purring at that time, then I knew this [*purring*] must mean purring [in Chinese]."
	2.2. Using images	-	Used the images of the video to support understanding of the TW	"At that time, I guessed it [*barneys*] should be… because the image showed the house, so I thought it should be the name of the residence or that kind of building."

Table 3: (continued)

Strategy categories	Subcategories	Further subcategories	Definitions	Example quote
	2.3. Using global understanding	–	Used the video plot or background knowledge to support understanding of the TW	"I noticed this word [*bizarre*] at that time. The context and plots before and after all indicated the situation was weird."
	2.4. Using local contextual cues	–	Used the word(s) that appeared just before and/or after the TW to support understanding of the TW	"Oh, this word [*sedated*] I guessed its meaning when I watched. Because the following is 'they are real', so the 'not' before the word should mean they are not. It has the meaning of virtual, or digital."
3. Using L1 translations	3.1. Using L1 matched	3.1.1. L2 triggered reference to L1	Noticed the L2 TW first, then checked its L1 translation	"Yes, this word [*sedated*] I didn't know, so I took a look at the L1 subtitles."
		3.1.2. L1 triggered reference to L2	Noticed the L1 translation of the TW first, then checked how to express the meaning in L2	"At that time, I was curious about how to express confiscation [in Chinese], so I took a look at the English word [*confiscated*]."
		3.1.3. Using L1 no sequence mentioned	Noticed both the L1 translation and L2 TW but not mentioned the particular sequence of awareness	"I heard '*hump*', and then I knew it means hump [in Chinese]."
	3.2. Using L1 not matched	3.2.1. L1 triggered other L2	Noticed the L1 translation of the TW first, then thought about other L2 words sharing similar meaning but not the TW	"[For TW *midwife*] Because I seemed to want to hear the pronunciation of 'nurse', but I didn't hear 'nurse'…"
		3.2.2. Mismatched L1 and L2	Failed to match the L1 translation with the L2 TW	"Oh I saw the Chinese subtitles. She said how to buffer stress [in Chinese], but I didn't know which word was '*buffer*'."
		3.2.3. Lack of time to check L1	Noticed the L2 TW but did not have time to check its L1 translation	"That [*foraging*], because the subtitles run fast, sometimes I may not have time to read the Chinese meaning."

Table 3: (continued)

Strategy categories	Subcategories	Further subcategories	Definitions	Example quote
4. Guessing without reported sources	4.1. Meaning fully guessed	–	Reported a guessed meaning with no specific source mentioned	"I heard it [*surrogate*] at that time, and then I knew it was, it was something like a surrogate mother [in Chinese]."
	4.2. Meaning partially guessed	–	Reported a partially guessed meaning with no specific source mentioned	"Hmm… something good, I didn't know the exact meaning, but I only knew it [*endearing*] means something nice."
	4.3. Meaning unsuccessfully guessed	–	Reported attempts to guess the meaning of the TW but not reported outcomes	"I don't know what it [*sedated*] means, and at that time I didn't know the meaning either, but I have attempted to guess, I think."
5. Other strategies	5.1. Pretest impact	–	Mentioned remembered having seen the TW in the pretest during their viewing	"Because I have seen this word [*midwife*] last time when completing the test, and then this time when I saw this word, I paid attention to it in particular."
	5.2. Dictionary use	–	Attempted to remember the TW and refer to dictionary afterwards	"I wanted to know its [*barneys*] meaning, and I even wanted to memorise it and to check the dictionary after."
	5.3. Visualising	–	Created mental images for the TW	"I saw it [*hump*], it didn't show the way it was humped [in Chinese], but I could image what it would look like."
6. No reported strategies	6.1. No reported meaning guessed	–	Noticed the unknown word but did not report attempts to guess or make form-meaning link for the TW	"I didn't know. I just felt I didn't know the meaning (of the TW *foraging*) at that time."
	6.2. Forgot thoughts	–	Forgot their thoughts when noticed the TW	"I saw this word [*twirls*]. But I don't remember my thoughts now, yes."

TWs. The frequency with which each strategy was used in the three groups is summarised in Table 4. In general, for the six broad categories, categories 3. Using L1 translations (31.60 %), 2. Using context (27.51 %), and 6. No reported strategies (22.30 %) were the most frequent categories that emerged across groups. To be specific, the top five most frequently reported sub-strategies across groups were: 3.1. Using L1 matched (29.74 %), 6.1. No reported meaning guessed (20.82 %), 2.2. Using images (10.04 %), 2.4. Using local contextual cues (8.55 %), and 2.3. Using global understanding (5.95 %).

4.3 RQ4 – vocabulary processing strategies by subtitling groups

Table 4 shows that all three subtitling groups reported their use of 1. Word feature analysis, 2. Using context, 5. Other strategies, and 6. No reported strategies to process unknown TWs. However, strategy 3. Using L1 translations only applied to L1 and bilingual subtitles groups, and strategy 4. Guessing without reported strategies was only reported by the captions group. To further compare the strategies used across different subtitling conditions, three Kruskal–Wallis Tests were run to check the potential differences in three general strategy categories: 1. Word feature analysis, 2. Using context, and 6. No reported strategies. The other three strategies were not analysed using Kruskal–Wallis Tests because two (3. Using L1 translations and 4. Guessing without reported sources) were not applicable for all three conditions, and 5. Other strategies only included very few cases ($n = 6$). Therefore, the results of these three strategies are only compared descriptively.

As can be observed in Table 4, for strategy 1. Word feature analysis, the captions group reported more cases of using word features than other groups. However, no statistically significant differences were revealed among the three subtitling groups, H (2, $n = 45$) = 4.12, $p = 0.13$. In terms of strategy 2. Using context, the reported cases were 2–3 times more in the captions group (43.22 %) than in L1 (17.24 %) and bilingual subtitles (13.98 %) groups. This group difference was further confirmed in Kruskal–Wallis Test, H (2, $n = 45$) = 11.42, $p = 0.003$. Follow-up pairwise comparisons indicated that participants in the captions group used context significantly more frequently than L1 group, U ($N_{Captions}$ = 15, N_{L1} = 15) = 53.00, $z = -2.51$, $p = 0.012$, $r = 0.46$, with a medium effect size (Cohen 1988), and than the bilingual subtitles group, U ($N_{Captions}$ = 15, $N_{Bilingual}$ = 15) = 35.50, $z = -3.24$, $p = 0.001$, $r = 0.59$, with a large effect size (Cohen 1988). This difference, as revealed in Table 4, was largely attributed to the different use of global understanding and local contextual cues. However, no significant difference in using context was revealed between L1 and bilingual subtitles groups, U (N_{L1} = 15, $N_{Bilingual}$ = 15) = 112.00, $z = -0.02$, $p = 0.98$, $r = 0.004$.

Table 4: Frequency of vocabulary processing strategies (Level 2) with unknown target words in stimulated recalls by group.

Strategy categories	Subcategories	Captions (%)	L1 (%)	Bilingual (%)	Total (%)
1. Word feature analysis		13 (11.02)	2 (3.45)	7 (7.53)	22 (8.18)
	1.1. Analysing part of speech	3 (2.54)	1 (1.72)	1 (1.08)	5 (1.86)
	1.2. Analysing word-structure	5 (4.24)	0 (0)	0 (0)	5 (1.86)
	1.3. Form association	3 (2.54)	0 (0)	4 (4.30)	7 (2.6)
	1.4. Analysing word pronunciation	2 (1.69)	1 (1.72)	1 (1.08)	4 (1.49)
	1.5. Word usage	0 (0)	0 (0)	1 (1.08)	1 (0.37)
2. Using context		51 (43.22)	10 (17.24)	13 (13.98)	74 (27.51)
	2.1. Using auditory cues	2 (1.69)	3 (5.17)	3 (3.23)	8 (2.97)
	2.2. Using images	13 (11.02)	6 (10.34)	8 (8.60)	27 (10.04)
	2.3. Using global understanding	13 (11.02)	1 (1.72)	2 (2.15)	16 (5.95)
	2.4. Using local contextual cues	23 (19.49)	0 (0)	0 (0)	23 (8.55)
3. Using L1 translations		0 (0)	27 (46.55)	58 (62.37)	85 (31.60)
	3.1. Using L1 matched	0 (0)	25 (43.10)	55 (59.14)	80 (29.74)
	3.2. Using L1 not matched	0 (0)	2 (3.45)	3 (3.23)	5 (1.86)
4. Guessing without reported sources		22 (18.64)	0 (0)	0 (0)	22 (8.18)
	4.1. Meaning fully guessed	5 (4.24)	0 (0)	0 (0)	5 (1.86)
	4.2. Meaning partially guessed	9 (7.63)	0 (0)	0 (0)	9 (3.35)
	4.3. Meaning unsuccessfully guessed	8 (6.78)	0 (0)	0 (0)	8 (2.97)
5. Other strategies		1 (0.85)	3 (5.17)	2 (2.15)	6 (2.23)
	5.1. Pretest impact	0 (0)	2 (3.45)	1 (1.08)	3 (1.12)
	5.2. Want to use a dictionary	1 (0.85)	0 (0)	0 (0)	1 (0.37)
	5.3. Visualising	0 (0)	1 (1.72)	1 (1.08)	2 (0.74)
6. No reported strategies		31 (26.27)	16 (27.59)	13 (13.98)	60 (22.30)
	6.1. No reported meaning guessed	31 (26.27)	15 (25.86)	10 (10.75)	56 (20.82)
	6.2. Forgot thoughts	0 (0)	1 (1.72)	3 (3.23)	4 (1.49)
Total (%)		118 (100)	58 (100)	93 (100)	269 (100)

Table 5: Frequency and percentage of subcategories within the "3. Using L1 translations" category in L1 and bilingual subtitles groups.

Subcategories	Further subcategories	L1 (%)	Bilingual (%)
3.1. Using L1 matched		25 (43.10)	55 (59.14)
	3.1.1. L2 triggered reference to L1	7 (25.93)	8 (13.79)
	3.1.2. L1 triggered reference to L2	0 (0)	5 (8.62)
	3.1.3. Using L1 no sequence mentioned	18 (66.67)	42 (72.41)
3.2. Using L1 not matched		2 (3.45)	3 (3.23)
	3.2.1. L1 triggered other L2	2 (7.41)	0 (0)
	3.2.2. Mismatched L1 and L2	0 (0)	1 (1.72)
	3.2.3. Lack of time to check L1	0 (0)	2 (3.45)
Total		27 (100)	58 (100)

Regarding strategies 3. Using L1 translations and 4. Guessing without reported strategies, Table 4 shows that the captions group did not report any cases of using L1 translations; instead, participants using captions tended to directly guess the meaning of unknown words, unlike the other two groups where L1 translations were available. For the bilingual and L1 subtitles groups, using L1 translations was the most frequently used strategy, with the bilingual subtitles group reporting more cases (62.37 %) than the L1 subtitles group (46.55 %). Further analysis of the subcategories showed that, on most occasions, participants in the bilingual and L1 subtitles groups could successfully match L1 translations to unknown TWs during viewing, with more successful cases reported in the bilingual subtitles group. Analysis of the further subcategories revealed a difference between the bilingual and L1 subtitles groups. Table 5 shows that the L1 translations in bilingual subtitles allowed participants to check their corresponding L2 word forms, whereas the use of L1 subtitles might mislead participants to match another L2 word to the L1 translations due to the lack of L2 written texts.

As can be seen from Table 4, participants had a relatively high rate of reporting no strategies for the noticed unknown TWs (i.e., 6. No reported strategies), likely due to the incidental nature of the viewing activity. This category seemed to be reported less frequently in the bilingual group compared to other groups. However, no statistically significant differences were revealed among the three groups, H $(2, n = 45) = 1.60, p = 0.45$.

5 Discussion

The present study aimed at providing a comprehensive understanding of how L2 learners process unknown vocabulary during subtitled viewing in different

subtitling conditions. The results are now discussed in relation to the two main aspects of processing examined, i.e., participants' reported awareness of the unknown TWs and vocabulary processing strategies.

5.1 Awareness of the unknown TWs

RQ1 and RQ2 aimed at exploring participants' awareness of the unknown vocabulary during subtitled viewing and potential group differences. Overall, participants reported awareness of 40.72 % of the unknown vocabulary. This awareness rate was slightly higher than the 32.6 % cases reported in reading studies (e.g., Godfroid and Schmidtke 2013) but lower than the 75.6 % cases reported in listening studies (e.g., Cai and Lee 2012). This indicates that, compared to listening-only, L2 learners were less likely to be aware of the unknown TWs during subtitled viewing where dynamic images were also available. The low awareness rate also attests to the nature of incidental learning. When comprehension is the focus of the viewing activity, L2 learners are more likely to ignore unknown words with their main focus on comprehending the L2 input.

The between group comparisons revealed that participants using captions were aware of significantly more unknown TWs than those using L1 subtitles. This result corroborated previous findings showing that captions were more beneficial than L1 subtitles for learning word forms (e.g., Peters 2019; Peters et al. 2016; Wang and Pellicer-Sánchez 2022), suggesting the benefits of presenting unknown words in both written and auditory forms to increase L2 learners' awareness of the words (Vanderplank 2019). This finding also supports the claims that the use of L1 subtitles may prevent learners from concentrating on spoken L2 and so bypass the L2 spoken form of unknown words (e.g., Peters 2019). The bilingual subtitles group reported a relatively lower awareness rate than the captions group, suggesting that the presence of L1 translations might potentially distract learners' attention away from L2 unknown words. This also concurs with the eye-tracking findings reported by Wang and Pellicer-Sánchez (2022), which revealed significant shorter time on L2 unknown TWs than corresponding L1 translations when using bilingual subtitles. However, it should be noted that no statistically significant differences emerged between bilingual subtitles and other conditions.

5.2 Vocabulary processing strategies

In response to RQ3, stimulated recall data showed that participants used various types of strategies to process unknown words that appeared only once in the video,

which echoes Rott's (2000) finding in the context of learning from reading. Similar to previous studies (e.g., Cai and Lee 2012; Hu and Nassaji 2012; Rott 2000), strategies in the present research were also found to be used alone as well as in combination to construct participants' knowledge of unknown words.

The types of strategies generated in the present study were similar to previous research exploring L2 learners' vocabulary inferencing or processing strategies during viewing, reading, and listening. The majority of strategies reported by Sydorenko (2010) were also recorded in the present study, such as using images, using local contextual cues, and analysing word features. Only one strategy (i.e., recognising words that are similar to L1) reported by Sydorenko was not found in the present research. This can be attributed to the lack of cognates between participants' L1 (Chinese) and L2 (English) in the present research. Some strategies that emerged in the present study were not recorded by Sydorenko (2010), such as using L1 translations and guessing without reported sources. These can be attributed to the lack of L1 translations in Sydorenko's (2010) research where only captions were examined. Besides, Sydorenko (2010) only used a questionnaire to investigate participants' strategy use, which might have only revealed some general strategies. In contrast, the present research examined strategy use for each individual word, and thus participants reported more specific and detailed strategies and also mentioned many cases where they did not use any particular strategies, which may be a better representation of incidental vocabulary learning from viewing.

In the current study, the most frequently reported strategy was using L1 translations (in the L1 and bilingual subtitles conditions), supporting those reading studies showing that when L1 was available, L2 learners frequently used L1 to aid their understanding of unknown words (e.g., Fraser 1999; Prichard and Atkins 2021; Rott 2005). This is expected since L1 translation is believed to be the most effective method to support understanding and help to build up an initial form-meaning link (Nation 2003; Schmitt 2008).

The second frequent category across groups was using context, which supports previous reading and listening research (e.g., Cai and Lee 2012; Fraser 1999; Lawson and Hogben 1996; Rott 2000, 2005; Vandergrift 2003). Notably, context in this study includes both written/auditory contextual cues and images, which is different from reading and listening research in which only written/verbal context is referred to. When images were presented, they became a strong competitor for verbal context in supporting learners' understanding of unknown words. This echoes Sydorenko's (2010) finding that using images was the most frequently used strategy to learn unknown words during captioned viewing. Moreover, Peters (2019) showed that unknown words with imagery support were more likely to be learned than those without imagery support. The present results would explain this finding by showing that when imagery support is available, learners do tend to use this processing strategy to support their learning.

The third most frequent category was reporting awareness of unknown words but not reporting guesses/thoughts concerning any aspects of the word. This could indicate that either participants did not use any strategies, or they did not remember. Reading and listening research has similarly reported that readers demonstrated awareness of TWs but did not infer the word meaning (e.g., Cai and Lee 2012; Fraser 1999; Rott 2000, 2005). These cases can be considered an exemplar of shallow processing according to the depth of processing hypothesis (Craik and Lockhart 1972), as words were only registered at the awareness level, but no further conscious effort was involved. Interestingly, similar cases were not frequently reported where vocabulary learning was the aim of reading (e.g., Fraser 1999). Its frequent occurrence in the present research also attests to the nature of incidental learning where the main aim was comprehension rather than language learning. This finding also helps to explain the inconsistent relationship between participants' attention to unknown words in incidental vocabulary learning and their learning gains in eye-tracking research (e.g., Montero Perez et al. 2015; Wang and Pellicer-Sánchez 2022), as increased attention to the word might not always indicate learners' active engagement but rather shallow processing.

5.3 Vocabulary processing strategies: between-group comparisons

In response to RQ4, the analyses revealed that the strategies used by bilingual and L1 subtitles groups differed significantly from the captions group in three main ways. Firstly, as expected, using L1 translations was the most frequently used strategy in the bilingual and L1 subtitles groups. Secondly, when L1 translations were not available, learners sometimes would resort to guessing the meaning of unknown words. Lastly, the captions group reported more frequent use of the context (including auditory cues, imagery, and verbal context) to process unknown TWs than the other groups, especially the use of verbal contextual cues.

These differences suggest that, participants using captions relied more on the verbal context or on other unmentioned sources to infer word meanings. Nevertheless, when L1 translations were also presented, participants would reasonably turn to L1 translations as a shortcut instead of using contextual cues. These findings also offer empirical evidence to explain why the use of bilingual subtitles seems to result in higher meaning recognition gains than captions, but not in meaning recall (e.g., Wang and Pellicer-Sánchez 2022). According to the depth of processing theory (Craik and Lockhart 1972), inferring meaning from context requires more mental effort which could lead to deeper processing and learning. However, it can be argued that having translations of L2 unknown words in bilingual subtitles may have

reduced learners' cognitive analysis of word meanings and led to shallower memory traces which were not enough to develop the ability to recall the meanings of newly learned words.

The similarities between the L1 and bilingual subtitles groups suggest that, when the L1 input is available, learners seem to use similar strategies by relying more on L1 translations. Importantly, when using bilingual subtitles, L1 translations might also serve as a trigger for learners to refer to L2 word forms, whereas L1 translations in L1 subtitles might have triggered participants' memory of other L2 words, resulting in mismatches between L1 translations and L2 TWs and potentially bypassing the auditory L2 input. This difference suggests that the presence of both written L1 and L2 in bilingual subtitles may encourage participants' use of the translations to process unknown words. This finding also echoes the comment from a participant in Li's (2016) study about processing unknown vocabulary using bilingual subtitles: "I can compare with the two lines of languages. I use L1 subtitles for getting the meaning and L2 subtitles for getting the words' spelling" (p. 195). When using bilingual subtitles, learners could simultaneously access the written forms of both L1 and L2, increasing their chances to successfully link auditory word forms with their written forms and L1 translations, which is more likely to facilitate a correct initial establishment of form-meaning connections for unknown words. Thus, compared to L1 subtitles, bilingual subtitles might have facilitated learners' engagement with L2 word forms and potentially led to an even greater reliance on L1 translations. This finding also corroborates the eye-tracking findings by Wang and Pellicer-Sánchez (2022) that participants using bilingual subtitles spent significantly longer time processing the L1 translations of unknown words than those using L1 subtitles.

6 Limitations

There are several limitations that need to be acknowledged. Firstly, this research only focused on intermediate-to-advanced-level L2 learners who were experienced in using on-screen text, including bilingual subtitles. Beginners or L2 learners with limited experience of using on-screen text might employ different vocabulary processing strategies. In addition, the inferential analyses were based on a limited sample size (N = 45); thus, the results of cross-group comparisons should be generalised with caution. The second limitation has to do with the operation of the stimulated recalls. As suggested by Gass and Mackey (2017), to combat the memory decay and reactivity issue, stimulated recall should be administered as soon as possible after the activity. However, the present study had posttests between the viewing activity and the stimulated recalls that were needed to address the research questions reported in Wang and Pellicer-Sánchez (2022). Consequently, there were a few

cases where participants failed to recall their processes of the unknown words. Moreover, the vocabulary posttests could have had an effect on participants' stimulated recall data. Future research should administer stimulated recalls immediately after participants' viewing or incorporate a design where some participants do not complete the vocabulary tests. Thirdly, the present research only focused on participants' initial processing of unknown words that appeared only once in the video. It would also be worth exploring how processing might differ throughout repeated encounters with unknown vocabulary. Additionally, the present study focuses on learners' processing of unknown words without distinguishing those that were actually learned after viewing; thus, it would be informative for future research to compare L2 learners' strategies used for those unknown words that were learned after viewing and those that were not learned. Based on our initial observation of the present data, there was no guarantee that a particular strategy would necessarily lead to successful learning. Systematic analysis is needed to properly investigate the relationship between strategy use and learning gains by taking into account word-related factors (e.g., word length, part of speech, imagery support). Lastly, future studies should also triangulate stimulated recall data with vocabulary learning gains and eye-tracking data to obtain a more comprehensive understanding of the learning processes and outcomes during viewing activities.

7 Conclusions

The present research was the first to collect qualitative data to investigate L2 learners' awareness and processing strategies of unknown words during viewing in different subtitling conditions. In line with the incidental nature of the learning setting examined here, participants were largely unaware of the unknown vocabulary, only reporting awareness for 40.72 % of the target vocabulary. Results showed that, when participants were aware of the unknown vocabulary, they reported having used a variety of strategies, and subtitling type led to differences in how the unknown words were processed. The most frequently reported strategies included using L1 translations (when available) and using context (including images and contextual cues). Interesting differences across subtitling conditions emerged. Captions supported participants' awareness of the form of unknown TWs when compared to L1 subtitles, while bilingual subtitles led to the use of more strategies to process the unknown vocabulary. Participants using captions relied more significantly on context, whereas participants using L1 and bilingual subtitles relied more on L1 translations. The frequent use of L1 translations could benefit the initial establishment of the form-meaning link for unknown words, but the less mental effort involved might result in shallower memory traces, which could explain the low meaning recall gains reported in earlier studies.

The present research has important pedagogical implications. This research provided qualitative findings to further confirm that even in the incidental learning context, L2 learners would still process some of the unknown words during subtitled viewing, which could potentially benefit their vocabulary learning. Different vocabulary processing strategies in viewing should be introduced to L2 learners because expanding the strategy repertoire and using more strategies helps to increase learners' vocabulary size (Gu 2020). L1 translations can be suggested to help establish the correct initial form-meaning connection for unknown words. Meanwhile, the use of images, sounds, and context to infer word meaning should also be encouraged, as the activation of both verbal and imagery systems could provide a stronger connection for learners' information processing and augment information recall (Paivio 1986). Language learners and classroom practitioners should also choose subtitling types based on different learning needs. Captions can be recommended when the aim is to focus on word forms. Bilingual subtitles can be used to facilitate the initial establishment of the form-meaning link for unknown words as they allow learners to match the word forms with correct meanings, which seems to be more advantageous than L1 subtitles. Notably, there were many cases where participants were not aware of the unknown vocabulary or did not actively engage with those words. Therefore, when the aim is to facilitate learners' vocabulary learning, classroom practitioners could consider applying methods to increase learners' awareness of the novel words (e.g., pre-teaching the unknown words, adding textual enhancement) and encourage their active engagement with unknown words to maximise the benefits of viewing for vocabulary learning.

References

BBC. 2013. Animal odd couples. Available at: https://www.bbc.co.uk/programmes/b03jp46s.

BBC. 2019. Subtitle guidelines (Version1.1.7). Available at: https://bbc.github.io/subtitle-guidelines/

Bilibili. n.d. 奇特的动物伙伴 1 [Odd animal couples 1]. Available at: https://www.bilibili.com/video/av21620515/.

BNC Consortium. 2007. The British National Corpus, XML edn. Oxford Text Archive. Available at: http://hdl.handle.net/20.500.12024/2554.

Bryman, Alan. 2012. *Social research methods*, 4th edn. Oxford: Oxford University Press.

Cai, Wei. & Benny P. H. Lee. 2012. Processing unfamiliar words: Strategies, knowledge sources, and the relationship to text and word comprehension. *Canadian Journal of Applied Linguistics* 15(1). 122–145.

Cohen, Jacob. 1988. *Statistical power analysis for the behavioral sciences*, 2nd edn. London: Routledge.

Corel Corporation. 2018. Corel VideoStudio Pro 2018 [Computer software]. https://www.videostudiopro.com/en/pages/videostudio-2018/.

Craik, Fergus. I. M. & Robert S. Lockhart. 1972. Levels of processing: A framework for memory research. *Journal of Verbal Learning & Verbal Behavior* 11(6). 671–684.

Ericsson, K. Anders & Herbert A. Simon. 1993. *Protocol analysis: Verbal reports as data*. Cambridge, MA: The MIT Press.

Fraser, Carol A. 1999. Lexical processing strategy use and vocabulary learning through reading. *Studies in Second Language Acquisition* 21(2). 225–241.

Gass, Susan M. & Alison Mackey. 2017. *Stimulated recall methodology in applied linguistics and L2 research*, 2nd edn. New York: Routledge.

Godfroid, Aline & Jens Schmidtke. 2013. What do eye movements tell us about awareness? A triangulation of eye-movement data, verbal reports and vocabulary learning scores. In Joara Martin Bergsleithner, Sylvia Nagem Frota & Jim Kei Yoshioka (eds.), *Noticing and second language acquisition: Studies in honor of Richard Schmidt*, 183–205. Honolulu: University of Hawaii, National Foreign Language Resource Center.

Godfroid, Aline. 2020. *Eye tracking in second language acquisition and bilingualism: A research synthesis and methodological guide*. New York: Routledge.

Gu, Peter Yongqi. 2020. Strategies for learning vocabulary. In Stuart Webb (ed.), *The Routledge handbook of vocabulary studies*. New York: Routledge.

Hu, Hsueh-chao Marcella & Hossein Nassaji. 2012. Ease of inferencing, learner inferential strategies, and their relationship with the retention of word meanings inferred from context. *Canadian Modern Language Review* 68(1). 54–77.

IELTS. (n.d.). IELTS in CEFR scale. Available at: https://www.ielts.org/about-ielts/ielts-in-cefr-scale.

Koolstra, Cees M. & Johannes W. J. Beentjes. 1999. Children's vocabulary acquisition in a foreign language through watching subtitled television programs at home. *Educational Technology Research & Development* 47(1). 51–60.

Lawson, Michael J. & Donald Hogben. 1996. The vocabulary-learning strategies of foreign-language students. *Language Learning* 46(1). 101–135.

Li, Mingyue. 2016. *Investigation into the differential effects of subtitles (first language, second language, and bilingual) on second language vocabulary acquisition*. Edinburgh: The University of Edinburgh PhD Dissertation.

Li, Mingyue & Mairin Hennebry-Leung. 2022. Effects of monolingual and bilingual subtitles on L2 vocabulary acquisition. *International Review of Applied Linguistics in Language Teaching*. https://doi.org/10.1515/iral-2022-0034, In press.

Lwo, Laurence & Michelle Chia-Tzu Lin. 2012. The effects of captions in teenagers' multimedia L2 learning. *ReCALL* 24(02). 188–208.

Mayer, Richard. 2009. *Multimedia learning*, 2nd edn. Cambridge: Cambridge University Press.

Montero Perez, Maribel. 2022. Second or foreign language learning through watching audio-visual input and the role of on-screen text. *Language Teaching* 55(2). 163–192.

Montero Perez, Maribel, Elke Peters & Piet Desmet. 2015. Enhancing vocabulary learning through captioned video: An eye-tracking study. *The Modern Language Journal* 99(2). 308–328.

Montero Perez, Maribel, Wim Van Den Noortgate & Piet Desmet. 2013. Captioned video for L2 listening and vocabulary learning: A meta-analysis. *System* 41(3). 720–739.

Nassaji, Hossein. 2003. L2 vocabulary learning from context: Strategies, knowledge sources, and their relationship with success in L2 lexical inferencing. *Tesol Quarterly* 37(4). 645–670.

Nation, Paul. 2001. *Learning vocabulary in another language*. Cambridge: Cambridge University Press.

Nation, Paul. 2003. The role of first language in foreign language learning. *The Asian EFL Journal* 5(2). 1–8.

Nation, Paul & Alex Heatley. 2002. Range: A program for the analysis of vocabulary in texts (Version 3) [Computer software]. Available at: https://www.lextutor.ca/cgi-bin/range/texts/index.pl.

Paivio, Allan. 1986. *Mental representations: A dual coding approach*. Oxford: Oxford University Press.

Peters, Elke, Eva Heynen & Eva Puimège. 2016. Learning vocabulary through audiovisual input: The differential effect of L1 subtitles and captions. *System* 63. 134–148.

Peters, Elke & Stuart Webb. 2018. Incidental vocabulary acquisition through viewing L2 television and factors that affect learning. *Studies in Second Language Acquisition* 40(3). 551–577.

Peters, Elke. 2019. The effect of imagery and on-screen text on foreign language vocabulary learning from audiovisual input. *Tesol Quarterly* 53(4). 1008–1032.

PortableSoft. 2012. SrtEdit (Version 6.3) [Computer software]. Available at: http://www.portablesoft.org/srtedit-portable/.

Prichard, Caleb & Andrew Atkins. 2021. Evaluating the vocabulary coping strategies of L2 readers: An eye tracking study. *Tesol Quarterly* 55(2). 593–620.

Puimège, Eva, Maribel Montero Perez & Elke Peters. 2023. Promoting L2 acquisition of multiword units through textually enhanced audiovisual input: An eye-tracking study. *Second Language Research* 39(2). 471–492.

Pujadas, Geòrgia & Carmen Muñoz. 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.

QSR International Pty Ltd. 2018. NVivo (Version 12). Available at: https://www.qsrinternational.com/nvivo-qualitative-data-analysis-software/home.

Rott, Susanne. 2000. Relationships between the process of reading, word inferencing, and incidental word acquisition, in assigning meaning to form. In James F. Lee & Albert Valdman (eds.), *Form and meaning: Multiple perspectives*, 255–282. Boston: Heinle & Heinle Publishers.

Rott, Susanne. 2005. Processing glosses: A qualitative exploration of how form-meaning connections are established and strengthened. *Reading in a Foreign Language* 17(2). 95–124.

Sanchez, Hugo Santiago & Trevor Grimshaw. 2020. Stimulated recall. In Jim McKinley & Heath Rose (eds.), *The Routledge handbook of research methods in applied linguistics*. New York: Routledge.

Schmitt, Norbert. 1997. Vocabulary learning strategies. In Norbert Schmitt & Michael McCarthy (eds.), *Vocabulary: Description, acquisition, and pedagogy*. Cambridge: Cambridge University Press.

Schmitt, Norbert, Diane Schmitt & Caroline Clapham. 2001. Developing and exploring the behaviour of two new versions of the Vocabulary Levels Test. *Language Testing* 18(1). 55–88.

Schmitt, Norbert. 2008. Review article: Instructed second language vocabulary learning. *Language Teaching Research* 12(3). 329–363.

Selvi, Ali Fuad. 2020. Qualitative content analysis. In Jim McKinley & Heath Rose (eds.), *The Routledge handbook of research methods in applied linguistics*, 440–452. New York: Routledge.

SR Research. 2016. Eyelink 1000 plus [Apparatus and software]. Mississauga, Ontario, Canada: SR Research. Available at: https://www.sr-research.com/eyelink1000plus.html.

SR Research Ltd. 2018. EyeLink Data Viewer (Version 3.1.246) [Apparatus and software]. Mississauga, Ontario, Canada: SR Research. Available at: https://www.sr-research.com/data-viewer/.

Sydorenko, Tetyana. 2010. Modality of input and vocabulary acquisition. *Language Learning & Technology* 14(2). 50–73.

Vandergrift, Larry. 2003. Orchestrating strategy use: Toward a model of the skilled second language listener. *Language Learning* 53(3). 463–496.

Vanderplank, Robert. 2010. 'Deja Vu'? A decade of research on language laboratories, television and video in language learning. *Language Teaching* 43(1). 1–37.

Vanderplank, Robert. 2019. 'Gist watching can only take you so far': Attitudes, strategies and changes in behaviour in watching films with captions. *The Language Learning Journal* 47(4). 407–423.

Wang, Andi & Ana Pellicer-Sánchez. 2022. Incidental vocabulary learning from bilingual subtitled viewing: An eye-tracking study. *Language Learning* 72(3). 765–805.

Wang, Andi & Ana Pellicer-Sánchez. 2023. Examining the effectiveness of bilingual subtitles for comprehension: An eye-tracking study. *Studies in Second Language Acquisition* 45(4). 882–905.

Webb, Stuart. 2020. Introduction. In Stuart Webb (ed.), *The Routledge handbook of vocabulary studies*, 1–12. New York: Routledge.