The Impact of Subtitling Modalities in Streaming Drama Series on EFL Vocabulary Acquisition

Wenhua Hsu*

***Corresponding author's email: whh@isu.edu.tw**

I-Shou University, Taiwan

Received: June 13, 2025 *Revised: November 6, 2025* *Accepted: December 4, 2025*

Abstract

This study investigated the effects of various subtitling modalities on vocabulary acquisition among English as a Foreign Language (EFL) learners in the context of binge-watching drama series. Four subtitling conditions were examined: (1) bimodal subtitling (L2 audio with L2 captions), (2) standard subtitling (L2 audio with L1 subtitles), (3) reversed subtitling (L1 audio with L2 subtitles), and (4) L3 audio with L2 subtitles. Eight groups of Taiwanese university students, divided into two proficiency levels, watched four drama series over a four-month period, rotating through each modality. Participants were instructed to avoid using dictionaries, and vocabulary tests were administered without providing grade incentives. The results revealed significant differences in vocabulary gains across subtitling modalities and proficiency levels. Upper-intermediate learners acquired an average of 20.89 out of 28 target words, while low-intermediate learners averaged 17.12 words. L3 audio with L2 subtitles yielded the highest gains, followed by bimodal and reversed subtitling, with standard subtitling resulting in the lowest performance. These findings suggest that EFL learners may also benefit from watching non-English-language dramas with English subtitles, expanding vocabulary through repeated exposure. English subtitles, analogous to graded readers in extensive reading programs, may offer valuable lexical input, regardless of the original language of the audiovisual content.

Keywords: bimodal subtitling, standard subtitling, reversed subtitling, L3 audio with L2 subtitles

Introduction

The rapid rise of over-the-top (OTT) streaming platforms has significantly changed how people watch television (TV) series. For Generation Y and Z, there has been a notable shift from traditional appointment viewing to streaming platform consumption. Appointment viewing refers to watching TV programs on a fixed schedule—typically once a day or once a week—while sitting on a couch in front of a television set. In contrast, streaming platforms accessible via mobile devices with Internet connectivity allow viewers to watch content anytime and anywhere, facilitating binge-watching. According to the Cambridge Dictionary, 'binge-watch' means watching several episodes of a program or TV series consecutively. This new viewing behavior has been described as watching two to six episodes in one sitting, based on a survey of nearly 1,500 U.S. adults who stream TV shows at least once a week (Netflix, 2014).

Streaming services have also contributed to the rise of binge-watching on college campuses in Taiwan, where English as a Foreign Language (EFL) is a required course for non-English majors. In Freshman English classes, the researcher-teacher often overhears her students discussing the drama series they have recently binge-watched. However, in private conversations, all of the students reported choosing Mandarin subtitles—their native language—when viewing, regardless of the language spoken in the videos. As a result, the researcher has expressed concern about the potential impact on English vocabulary development if students were to shift from using Mandarin subtitles to English ones.

Subtitling is an audiovisual translation modality. Subtitles are lines of text displayed at the bottom of the screen, translating spoken dialogues into the audience's first language, which differs from the language used in the source video (Nesirli, 2025). On OTT platforms targeting global audiences, English is often the primary language used for content translation. As a pivot translation, English subtitles frequently serve as the source text for translation into many other languages (Locher & Messerli, 2020; Pedersen, 2019).

Subtitles are on-screen texts in the viewer's first language (L1) as opposed to L2 captions in L2 videos (Markham & Peter, 2003). Talaván Zanón (2006) categorized subtitling modalities into three types: bimodal, standard, and reversed subtitling. Bimodal subtitles refer to on-screen text when the audio and subtitles are in the same foreign language (L2 audio with L2 captions). Standard subtitling involves on-screen text in the viewer's native language, paired with a video in a foreign language (L2 audio with L1 subtitles). In contrast, reversed subtitling (L1 audio with L2 subtitles) involves a scenario where the video is in the viewer's native language, and the subtitles are in a foreign language. Apart from these two common configurations (L1 subtitles and L2 captions in L2 videos), the current research context includes subtitles in the viewer's L2 (English) in L3 (Korean) videos. This setup may compel Taiwanese students with limited knowledge of the L3 to rely on English on-screen text when Mandarin subtitles are not used. In this situation, English subtitles in L3 dramas can be compared to graded readers

that English teachers, like ours, often encourage their EFL students to read as part of an extensive reading scheme.

Given the binge-watching phenomenon, this research examines the impact of various subtitling modalities on the lexical acquisition of EFL learners. Specifically, it examines L2 audio with L2 captions, L2 audio with L1 subtitles, L1 audio with L2 subtitles, and L3 audio with L2 subtitles, representing different audiovisual channels of English (L2) exposure. The research question guiding the investigation is as follows: *Is there a significant difference in vocabulary gains when EFL learners are exposed to different subtitling modalities while watching a drama series?*

Literature Review

Extensive Reading, Repetitions, and Incidental Learning

Vocabulary learning is a gradual process. In view of this, some vocabulary researchers promote extensive reading (Day & Bamford, 1998; Grabe, 2009; Krashen, 2004, 2011). For extensive reading to be effective, learners need to read in quantities large enough to have sufficient encounters with unfamiliar words so as to eventually internalize how they are used and how they collocate with the words around them (Nation, 2015).

Previous research on incidental learning of words has shown that learning rarely occurs after a single encounter (Horst, Cobb & Meara, 1998). As the number of encounters with an unknown word increases, the potential of learning the word increases, as repeated occurrences stand a favorable opportunity of increasing the word's prominence and hence the learning thereof. However, there is no particular repetition threshold that assures word acquisition, since the effort needed to learn the full meaning of a word depends on the word itself (e.g., multiple meanings), the context (the amount of information provided to infer meanings) (Webb, 2008), and learners' proficiency (Zahar, Cobb & Spada, 2001). Based on prior results ranging from two to twenty repetitions, Nation (2014) took a middle ground and chose 12 repetitions as the cutoff to measure the amounts of input needed for incidental learning of most of the first 9000-word families. This research also adopted a frequency threshold of 12 occurrences when selecting target words appearing in the English subtitles of the four drama series.

Past Studies on Subtitling Modalities

Research on subtitling modalities has consistently demonstrated that the type of subtitle displayed during video viewing influences comprehension and vocabulary learning among EFL learners. Studies have primarily compared standard subtitling (L2 audio with L1 subtitles), bimodal subtitling (L2 audio with L2 captions), and reversed subtitling (L1 audio with L2 subtitles).

Using American TV shows streamed on Netflix, Dizon (2016) examined the perceptions of intermediate Japanese EFL learners regarding three subtitling modes—L1 subtitles, L2 captions, and

no subtitles—and found that L2 captions were preferred over the other options. Subsequent research (Dizon, 2018, 2021) similarly indicated that L2 captions enhanced comprehension and vocabulary acquisition, suggesting that EFL learners value L2 textual input as an aid for language learning. These findings highlight the potential of captioned streaming content to promote informal vocabulary learning.

Early research on reversed subtitling revealed pedagogical advantages. Lambert et al. (1981) found that reversed subtitling (L1 audio with L2 subtitles) most effectively improved comprehension and contextual meaning among French-immersion students, as L1 audio facilitated L2 text processing. Holobow et al. (1984) also confirmed that sustained exposure to reversed subtitling yielded the highest comprehension and contextual-meaning scores, concluding that beginning L2 learners benefit when the L1 is available in at least one modality. Similarly, Talaván Zanón (2006, 2007) reported that reversed subtitling supports vocabulary and grammar learning among less advanced learners.

Drawing on dual coding theory (Paivio, 1971, 1986), Danan (1992, 2004) explained that reversed subtitling allows learners to associate visual imagery, L1 audio, and L2 text, thereby strengthening verbal and visual memory traces. Comprehensible L1 input combined with L2 textual information promotes deeper processing and facilitates incidental vocabulary learning. In contrast, standard subtitling may overload working memory by requiring learners to process both L2 speech and L1 text, which differ syntactically.

Several comparative studies have found that reversed subtitling is particularly effective for improving comprehension and vocabulary recall. Bairstow and Lavaur (2011) observed that reversed subtitling outperformed both bimodal subtitling and standard subtitling in vocabulary recall and comprehension. Similarly, Fazilatfar et al. (2011) reported that reversed subtitling enhanced word-form recognition and meaning translation more than standard subtitling or no subtitling. Positive learner perceptions of reversed subtitling were also noted by Čepon (2011), who found that participants believed the approach supported vocabulary development.

More recent research has produced mixed findings. Gorijan (2014) found that reversed subtitling was the most effective modality for incidental vocabulary learning among Iranian EFL students, followed by bimodal and standard subtitling. Conversely, Kwarteng (2023) reported that standard subtitling led to greater vocabulary improvement than bimodal subtitling, suggesting that factors such as proficiency level, video type, and linguistic distance may moderate the effects of subtitling. Through TOEIC listening and reading tests administered to 108 Korean EFL learners after ten weeks of exposure to subtitled materials, Kim (2020) observed differential benefits across receptive skills: reversed subtitling improved reading, whereas standard subtitling enhanced listening comprehension.

A large-scale meta-analysis by Reynolds et al. (2022), which reviewed 139 prior studies, confirmed that both captioned and subtitled videos have a positive influence on vocabulary acquisition. The authors emphasized that subtitling modality, language pairing, and learner proficiency level are critical variables shaping learning outcomes.

Despite robust evidence supporting the value of subtitled input, several gaps remain. Most existing studies have been conducted in controlled classroom settings using short video clips rather than authentic viewing experiences. Moreover, few studies have examined how different subtitling modalities influence vocabulary gains during binge-watching—a common mode of "serial viewing" on over-the-top (OTT) platforms (Rubenking & Bracken, 2021, p. 100356). However, Hsu (2023a, 2023b) drew parallels between extensive reading of graded readers and the voluminous reading of English subtitles prompted by K-drama fever and binge-watching, highlighting the potential for vocabulary development under the L3 audio with L2 subtitling modality. Addressing these gaps, the present study investigates whether significant differences in vocabulary gains emerge among EFL learners exposed to standard, bimodal, and reversed subtitling while binge-watching a drama series. By situating subtitled viewing within an authentic, sustained-exposure context, this study aims to clarify the comparative effectiveness of subtitling modalities in promoting incidental vocabulary learning.

Methodology

The Setting and the Participants

As mentioned earlier, Freshman English is a required course, and matriculating students are placed into classes at different proficiency levels based on their English scores from the college entrance exam. The researcher-teacher taught five Freshman English classes at the low-intermediate and upper-intermediate levels, with approximately 50 students in each class. In these five classes, 104 students volunteered to participate in this study and were compensated with a four-month Netflix subscription. The updated Vocabulary Levels Test (VLT) was administered in the classroom to measure their vocabulary size. According to Webb et al. (2017), the updated VLT consists of five levels, totaling 150 test items, with 30 items drawn from each of the BNC/COCA 1st to 5th 1000-word families. Based on the VLT scores, half of the participants had a vocabulary size within the 2,000–3,000-word family range (equivalent to the CEFR B1 level), while the other half fell within the 4,000–5,000-word family range (B2 level) (Milton, 2010; Nation, 2013). The 52 low-intermediate and 52 upper-intermediate students were divided into four groups of 13 participants each, totaling eight groups. Levene's tests for homogeneity of variances confirmed that the population variances among the four groups at both proficiency levels were approximately equal ($p > .05$ for both).

All participants were instructed to watch four designated dramas (American TV series *Emily in Paris* and *My Life with the Walter Boys*, the Mandarin drama *At the Moment*, and the Korean drama *Twenty-Five Twenty-One*), each to be completed within one month during their spare time. The four assigned drama series were selected based on their romance or coming-of-age themes, which are generally favored by college students, as well as their popularity in OTT platform ratings at the time of the experiment.

To provide an impression of the English subtitles the students were exposed to, Table 1 presents four excerpts in SRT format from Episode 1 of the English-subtitled American TV series *Emily in Paris* and *My Life with the Walter Boys*, the Mandarin drama *At the Moment*, and the Korean drama *Twenty-Five Twenty-One*. English subtitles in SubRip Text (SRT) format were downloaded online for research purposes. These files are plain text documents that include the sequence of subtitles along with their start and end time codes. Each scene contains one or two lines of subtitles, with a maximum of 7 to 8 words per line.

Table 1
Subtitle excerpts in SRT format

American TV series ***Emily in Paris*** **Season 2, Episode 1**	**American TV series** ***My Life with the Walter Boys*** **Season 1, Episode 1**	**Mandarin drama** ***At the Moment*** **Episode 1**	**Korean drama** ***Twenty-five Twenty-one*** **Episode 1**
15 00:01:36,638 --> 00:01:39,724 Are you okay? You look like you just lost a follower.	19 00:00:59,434 --> 00:01:02,771 I hope so. The photo booth alone has raised over a thousand dollars	24 00:02:36,166 --> 00:02:38,333 Don't forget your dream just because you failed."	15 00:00:41,176 --> 00:00:44,054 Just show them how much you love ballet.
16 00:01:39,808 --> 00:01:42,018 No, but maybe a friend.	20 00:01:02,854 --> 00:01:04,105 for charity so far.	25 00:02:38,416 --> 00:02:40,375 This quote is very important to me.	16 00:00:44,137 --> 00:00:45,764 Don't think about doing well.
17 00:01:42,102 --> 00:01:43,353 Camille.	21 00:01:04,689 --> 00:01:06,649 I'm gonna run for Events chair next semester.	26 00:02:40,458 --> 00:02:43,041 It reminds me of the courage to pursue my dream...	17 00:00:45,847 --> 00:00:48,516 How can I not think about that? I'm here to win.
18 00:01:43,436 --> 00:01:46,523 I haven't spoken to her since Gabriel told me he was staying in Paris,	22 00:01:06,733 --> 00:01:08,401 Jackie, we're freshmen.	27 00:02:43,125 --> 00:02:44,166 Cut.	18 00:00:48,600 --> 00:00:50,393 Winning doesn't matter.
19 00:01:46,606 --> 00:01:47,857 and now I have to go to Savoir	23 00:01:09,069 --> 00:01:11,029 Only juniors and seniors can be committee chairs.	28 00:02:44,666 --> 00:02:46,666 Can't you see how messy her hair is?	19 00:00:50,477 --> 00:00:53,396 That's unconvincing coming from you.
20 00:01:47,941 --> 00:01:50,068 to meet with her about her family champagne	24 00:01:11,112 --> 00:01:12,113	29 00:02:48,250 --> 00:02:49,250 Hairstylist.	20 00:01:04,366 --> 00:01:06,910 The 43rd Myungin Dance Contest

	We'll see.

Table 2 provides further details about the four drama series and the subtitling modalities used in this research.

Table 2
Four dramas under four subtitling modalities

Drama title	Genre	Language spoken	Subtitling language	Subtitling modality	Number of subtitled words and episodes
Emily in Paris S2	Romance, comedy	English (L2)	English (L2)	Bimodal subtitling	34,983 10 episodes
My Life with the Walter Boys S1	Romance, adolescence	English (L2)	Mandarin (L1)	Standard subtitling	51,907 10 episodes
At the Moment	Urban romance	Mandarin (L1)	English (L2)	Reversed subtitling	44,228 10 episodes
Twenty-five Twenty-one	Romance, coming of age	Korean (L3)	English (L2)	L3-to-L2 subtitling (See note)	101,429 16 episodes

Note: The L3 audio with L2 subtitling modality has not been named in the literature and is tentatively referred to as L3-to-L2 subtitling.

While watching the four designated dramas, the 104 participants were instructed to refrain from using dictionaries when encountering unfamiliar words. Furthermore, they were informed that their vocabulary test scores would not affect their course grades.

The Design of the Study

Table 3 outlines the study design. The eight groups of students across two proficiency levels took turns watching the same four designated dramas and therefore experiencing each of the four subtitling modalities: (1) bimodal subtitling (L2 audio with L2 captions), (2) standard subtitling (L2 audio with L1 subtitles), (3) reversed subtitling (L1 audio with L2 subtitles), and (4) L3 audio with L2 subtitles.

Table 3
The design of the study

Groups		Month/Round 1	Month/Round 2	Month/Round 3	Month/Round 4
Upper-intermediate (*n*=13)	Gp1	Drama 1 with Modality A	Drama 2 with Modality B	Drama 3 with Modality C	Drama 4 with Modality D
Low-intermediate (*n*=13)	Gp5				
Upper-intermediate (*n*=13)	Gp2	Drama 2 with Modality B	Drama 3 with Modality C	Drama 4 with Modality D	Drama 1 with Modality A
Low-intermediate (*n*=13)	Gp6				

Upper-intermediate (*n*=13) Low-intermediate (*n*=13)	Gp3 Gp7	Drama 3 with Modality C	Drama 4 with Modality D	Drama 1 with Modality A	Drama 2 with Modality B
Upper-intermediate (*n*=13) Low-intermediate (*n*=13)	Gp4 Gp8	Drama 4 with Modality D	Drama 1 with Modality A	Drama 2 with Modality B	Drama 3 with Modality C

Note: Modality A: L2 audio with L2 captions (bimodal subtitling); Modality B: L2 audio with L1 subtitles (standard subtitling); Modality C: L1 audio with L2 subtitles (reversed subtitling); Modality D: L3 audio with L2 subtitles
Drama 1: American TV series *Emily in Paris* S2; Drama 2: American TV series *My Life with the Walter Boys* S1; Drama 3: Mandarin drama *At the Moment*; Drama 4: Korean drama *Twenty-five Twenty-one*

One month before viewing the dramas, all participants took a vocabulary meaning-recognition pre-test to assess their knowledge of the target words. This one-month interval was intended to reduce memory retention of the target words. For each item on the test, the participants were asked to choose the correct Chinese meaning from a set of options. To minimize guessing, 'I don't know' was included as an option for each item (as below).

Test item #15. pathetic (A)可悲的 (B)卑鄙的 (C)糟糕的 (D)痛苦的 (E)不知道

Seven target words—each appearing 12 times (indicating an equivalent amount of visual English exposure; see Literature Review for repetitions and incidental learning)—were selected from each of the four dramas, totaling 28 target words. These words were chosen because they were beyond the 5th 1,000-word-family level and likely unfamiliar to the students. To prevent students from focusing solely on the target words, the 28 target words were mixed with 32 distractor words that did not appear in any of the four dramas.

At the end of each month (i.e., after viewing one designated drama), the participants took a post-test containing the seven target words from the drama they had just watched. After four months of viewing, 15 students–five who gained the most target words, five who performed in the middle, and five who gained the fewest target words–were interviewed about their perceptions of subtitling modalities across two languages while watching drama series. Finally, statistical tests were conducted using two-way analysis of covariance (ANCOVA) and independent t-tests in IBM SPSS Statistics 29.0.

Results and Discussion

The means and standard deviations of the vocabulary gain scores (the differences between pre-test and post-test scores) were presented in Table 4, along with the means adjusted by the covariate (pre-test scores) and the corresponding standard errors.

Table 4
Means, standard deviations, adjusted means, and standard errors of vocabulary gains across eight

groups

		Month/Round 1			**Month/Round 2**			**Month/Round 3**			**Month/Round 4**		
Level	Gp	Mode	M (SD)	M adj (SE)	Mode	M (SD)	M adj (SE)	Mode	M (SD)	M adj (SE)	Mode	M (SD)	M adj (SE)
Upper-intermediate *N*=13 for all groups	1	A	6.08 (0.760)	6.074 (0.195)	B	3.54 (0.519)	3.544 (0.171)	C	4.62 (0.768)	4.600 (0.203)	D	6.77 (0.439)	6.773 (0.199)
	2	B	3.62 (0.650)	3.621 (0.195)	C	4.54 (0.519)	4.544 (0.171)	D	6.62 (0.506)	6.621 (0.203)	A	6.23 (0.725)	6.224 (0.199)
	3	C	4.69 (0.751)	4.689 (0.195)	D	6.46 (0.519)	6.460 (0.171)	A	5.77 (0.832)	5.754 (0.203)	B	3.62 (0.768)	3.619 (0.199)
	4	D	6.69 (0.480)	6.707 (0.196)	A	6.15 (0.689)	6.152 (0.171)	B	3.54 (0.660)	3.544 (0.203)	C	4.62 (0.768)	4.598 (0.200)
Low-intermediate *N*=13 for all groups	5	A	5.08 (0.760)	5.074 (0.195)	B	2.54 (0.519)	2.537 (0.171)	C	3.62 (0.768)	3.621 (0.203)	D	6.00 (0.707)	6.004 (0.199)
	6	B	2.62 (0.650)	2.612 (0.195)	C	3.54 (0.519)	3.537 (0.171)	D	5.85 (0.801)	5.851 (0.203)	A	5.23 (0.725)	5.235 (0.199)
	7	C	3.69 (0.751)	3.689 (0.195)	D	5.69 (0.855)	5.691 (0.171)	A	4.77 (0.832)	4.774 (0.203)	B	2.62 (0.768)	2.619 (0.199)
	8	D	5.92 (0.760)	5.920 (0.195)	A	5.15 (0.689)	5.152 (0.171)	B	2.54 (0.660)	2.544 (0.203)	C	3.62 (0.768)	3.619 (0.199)

Notes: Gp=Group; M=The mean of the vocabulary gain scores (the difference between post-test and pre-test scores); M adj=Mean adjusted; SD=Standard Deviation; SE=Standard Error
The pre-test included 28 (4×7) target vocabulary items from four dramas. Pre-test 1 scores refer to the number of target words from Drama 1 that were answered correctly (full marks=7 points), while Pre-test 2 scores refer to the number of target words from Drama 2 that were answered correctly (full marks=7 points), and so on. Adjustments based on covariates, pre-test 1 mean=0.18; pre-test 2 mean=0.17; pre-test 3 mean=0.17; pre-test 4 mean=0.18

As shown in Table 4, eight groups of participants alternately received four subtitling modes while viewing the designated dramas. Each group was equally exposed to all four modalities, allowing for within-group comparisons (see Table 4, horizontally, for each group's data). Across both proficiency levels, the highest vocabulary gain scores consistently occurred under Modality D (L3 audio with L2 subtitles, specifically in this study, a Korean drama with English subtitles). The participants learned the most target words after 12 repetitions (M = 5.69–6.77; M_{adj} = 5.691–6.773 across groups). Modality A (L2 audio with L2 captions, or bimodal subtitling; in this case, an American TV series with English captions) produced the second-highest gains, followed by Modality C (L1 audio with L2 subtitles, or reversed subtitling; a Mandarin drama with English subtitles). Modality B (L2 audio with L1 subtitles, or standard subtitling; an American series with Mandarin subtitles) yielded the lowest vocabulary gains (M = 2.54–3.62; M_{adj} = 2.537–3.621).

Between-group comparisons confirmed the superior effect of Modality D. Regardless of proficiency level, participants who viewed dramas under this modality consistently learned more target words than those exposed to the other three subtitling types. In Month/Round 1, upper-intermediate Group 4 and low-intermediate Group 8—both receiving Modality D—achieved the highest mean scores (M = 6.69 and M = 5.92, respectively). Similar results were observed across subsequent rounds, with the groups assigned to Modality D outperforming all others.

Overall, Modality D (L3 audio with L2 subtitles) produced the greatest vocabulary gains, ranging from 6.45 to 6.77 words out of 7 target words (each appearing 12 times per drama series) for upper-intermediate groups and from 5.69 to 6.00 words for low-intermediate groups. The observed and adjusted means followed the same pattern: Modality A (bimodal subtitling) ranked second, Modality C (reversed subtitling) third, and Modality B (standard subtitling) last.

Additionally, upper-intermediate participants outperformed low-intermediate participants overall, irrespective of subtitling condition. An independent-samples t test showed that the upper-intermediate participants (M = 20.89, SD = 2.59) scored significantly higher than low-intermediate learners (M = 17.12, SD = 2.83), $t(102) = 15.39$, $p < .001$, MD = 3.62, SE = 0.24, 95% CI [3.15, 4.08].

To examine whether subtitling modality affects vocabulary gain and whether proficiency influences this effect, four rounds of two-way ANCOVA were conducted, using vocabulary pre-test scores as the covariate. The independent variables were the four subtitling modalities and two proficiency levels, while the dependent variable was the vocabulary gain scores.

The data contained no outliers, as assessed by box-plot inspection and by the absence of standardized residuals greater than ±3. The Shapiro-Wilks test indicated that the residuals were normally distributed ($p > .05$). There was homogeneity of regression slopes, as the interaction terms between the covariate (pre-test scores) and the independent variables (modality and proficiency) were not statistically significant ($p > .05$). Levene's test of equality of error variances confirmed that the assumption of homogeneity of variances was not violated ($p > .05$). The results of the four rounds of two-way ANCOVA are presented in Table 5.

Table 5

Four rounds of two-way ANCOVA results for vocabulary gain scores by four subtitling modalities and two proficiency levels

	Month/Round 1			**Month/Round 2**		
Source	F	Sig.	η_p^2	F	Sig.	η_p^2
Pre-test	0.415	.521	.004	0.304	.583	.003
Proficiency	47.038	< .001	.331	60.880	< .001	.391
Modality	106.650	< .001	.771	135.908	< .001	.811
Interaction	0.152	.928	.005	0.236	.871	.007
	Month/Round 3			**Month/Round 4**		
Source	F	Sig.	η_p^2	F	Sig.	η_p^2
Pre-test	2.021	.158	.021	0.591	.444	.006
Proficiency	42.018	< .001	.307	43.805	< .001	.316
Modality	93.194	< .001	.746	112.477	< .001	.780
Interaction	0.144	.933	.005	0.153	.927	.005

Notes: The Levene's test of equality of error variances was not significant ($p > .05$). The assumption of homogeneity of variances was not violated for all the groups (p = .851 for Round 1; p= .144 for Round 2; p= .652 for Round 3; p= .293 for Round 4).

Table 5 shows that subtitling modality had a statistically significant main effect on vocabulary learning across all four rounds [$F(3, 95) = 106.65$, $p < .001$, $\eta^2_p = .771$ in Round 1; $F(3, 95) = 135.91$, $p < .001$, $\eta^2_p = .811$ in Round 2; $F(3, 95) = 93.19$, $p < .001$, $\eta^2_p = .746$ in Round 3; $F(3, 95) = 112.48$, p

< .001, η^2_p = .780 in Round 4]. According to Cohen's (2013) benchmarks, these effect sizes indicate strong relationships between subtitling modality and vocabulary gains. Subtitling modality accounted for 77.1%, 81.1%, 74.6%, and 78.0% of the variance in vocabulary gain scores in Rounds 1 through 4, respectively, after controlling for pretest scores. These findings suggest that the subtitling modality used significantly influenced students' vocabulary learning while binge-watching drama series.

Proficiency level also had a statistically significant effect on vocabulary learning, though with smaller effect sizes [$F(1, 95) = 47.04$, $p < .001$, $\eta^2_p = .331$ in Round 1; $F(1, 95) = 60.88$, $p < .001$, $\eta^2_p = .391$ in Round 2; $F(1, 95) = 42.02$, $p < .001$, $\eta^2_p = .307$ in Round 3; $F(1, 95) = 43.81$, $p < .001$, $\eta^2_p = .316$ in Round 4].

No significant interaction was found between subtitling modality and proficiency [$F(3, 95) = 0.15$, $p = .928$ in Round 1; $F(3, 95) = 0.24$, $p = .871$ in Round 2; $F(3, 95) = 0.14$, $p = .933$ in Round 3; $F(3, 95) = 0.15$, $p = .927$ in Round 4]. This indicates that the influence of subtitling modality on vocabulary learning was consistent across proficiency levels.

Table 6

Pairwise comparisons and mean differences in vocabulary gains by four subtitling modalities controlling for pre-test

Month/Round 1					**Month/Round 2**				
	Upper-intermediate		Low-intermediate			Upper-intermediate		Low-intermediate	
Modality	*MD*	Sig.	*MD*	Sig.	Modality	*MD*	Sig.	*MD*	Sig.
A-B	2.453	.000	2.462	.000	A-B	2.608	.000	2.615	.000
A-C	1.385	.000	1.385	.000	A-C	1.608	.000	1.615	.000
A-D	-0.633	.004	-0.846	.003	A-D	-0.308	.097	-0.539	.084
B-C	-1.068	.000	-1.077	.000	B-C	-1.000	.000	-1.000	.000
B-D	-3.086	.001	-3.308	.000	B-D	-2.916	.003	-3.154	.000
C-D	-2.018	.001	-2.231	.000	C-D	-1.916	.003	-2.154	.000
Month/Round 3					**Month/Round 4**				
	Upper-intermediate		Low-intermediate			Upper-intermediate		Low-intermediate	
Modality	*MD*	Sig.	*MD*	Sig.	Modality	*MD*	Sig.	*MD*	Sig.
A-B	2.210	.000	2.230	.000	A-B	2.605	.000	2.616	.000
A-C	1.154	.000	1.153	.000	A-C	1.626	.000	1.616	.000
A-D	-0.867	.001	-1.077	.000	A-D	-0.549	.009	-0.769	.006
B-C	-1.056	.001	-1.077	.000	B-C	-0.979	.000	-1.000	.000
B-D	-3.077	.001	-3.307	.000	B-D	-3.154	.001	-3.385	.000
C-D	-2.021	.001	-2.230	.000	C-D	-2.175	.002	-2.385	.000

Notes: MD: mean difference; Modality A: L2 audio with L2 captions (bimodal subtitling); Modality B: L2 audio with L1 subtitles (standard subtitling); Modality C: L1 audio with L2 subtitles (reversed subtitling); Modality D: L3 audio with L2 subtitles Based on estimated marginal means. Adjustment for multiple comparisons: Boferroni. The mean difference is significant at the .05 level.

Subsequently, four rounds of post-hoc tests were conducted to evaluate pairwise differences among the adjusted means, using the Bonferroni correction to control for Type I error. The vocabulary gain scores were adjusted to account for initial differences (i.e., controlling for pre-test scores). Table 6 shows that there were significant differences in vocabulary gain scores between the upper-

intermediate group receiving Modality D and the other three upper-intermediate groups receiving Modalities A, B, and C; the same pattern was observed for the low-intermediate groups (all $p < .05$ except in Round 2, where the difference between Modalities A and D was not significant, $p > .05$). Ignoring the direction of the differences, the largest mean differences were consistently found between the Modality D and Modality B groups across both proficiency levels and all four rounds, followed by differences between the Modality A and Modality B groups. The smallest–but still significant–differences were observed between the Modality C and Modality B groups.

To sum up, regarding the effect of subtitling modality on vocabulary learning, Modality D (L3 audio with L2 subtitles) ranked highest, followed by Modality A (L2 audio with L2 captions, or bimodal subtitling), Modality C (L1 audio with L2 subtitles, or reversed subtitling), and finally Modality B (L2 audio with L1 subtitles, or standard subtitling), which ranked lowest.

The strong performance of Modality D (L3 audio with L2 subtitles) in promoting vocabulary learning through binge-watching drama series holds important implications for L2 learners. In this context, when Taiwanese college students choose English (L2) subtitles over Mandarin (L1) while watching a Korean (L3) drama, they rely on the subtitles to understand character dialogues because of their limited knowledge of Korean. This process resembles extensive reading, where reading for pleasure over time enables learners to repeatedly encounter unfamiliar words. For extensive reading to be effective, learners must engage with enough content to encounter new words frequently, allowing them to internalize word usage and collocations in context (Nation, 2015). The success of Modality D may also offer encouragement to English learners who enjoy binge-watching non-English-language dramas. As such, their viewing choices need not be limited to English-speaking series, as long as English subtitles are available. Similar to graded readers in extensive reading programs, English subtitles in this context can serve as a valuable resource for vocabulary acquisition.

The second-best performance by the bimodal subtitling groups (Modality A) in vocabulary gain scores may support the dual-coding hypothesis, which suggests that learning improves when related information is presented simultaneously through both auditory and visual channels, rather than through either channel alone. Proponents of the dual-coding hypothesis emphasize the integration of dually coded information into mental representations, arguing that this integration can reduce cognitive load on either the visual or auditory channel. This mode of delivery involves three elements: image, sound, and text. The rationale for presenting the same words in two formats (spoken and on-screen text) is that students can engage with the format that best suits their learning style. Those who learn better from spoken language can focus on character dialogues, while those who learn better from written input can rely on captions.

Contrary to past findings (Bairstow & Lavaur, 2011; Danan, 1992; Fazilatfar et al., 2011; Gorjian, 2014; Holobow et al., 1984; Lambert et al., 1981; Talaván Zanón, 2006, 2007), which suggested that reversed subtitling yields the greatest effect on vocabulary learning, the present results show that the reversed subtitling groups (Modality C: L1 audio with L2 subtitles) outperformed only

the standard subtitling groups (Modality B: L2 audio with L1 subtitles). The reason for this outcome is illuminated in the interview results. Nevertheless, regardless of the subtitling modality, learners appear to benefit from exposure to English, as reflected in the vocabulary gain scores–all of which were greater than zero under Modalities A, B, C, and D (see Table 4).

Interview Results

Fifteen students–selected from the top, middle, and bottom tiers of vocabulary gain scores–were interviewed to explore their experiences with binge-watching drama series. Thematic analysis of their responses identified salient themes: (1) adaptation to English subtitles, (2) plot engagement as a viewing motivator, (3) comprehension of Korean dramas with English subtitles, (4) engagement with English subtitles in Mandarin dramas, and (5) challenges associated with standard subtitling (L2 audio with L1 subtitles).

Theme 1: Adaptation to English Subtitles

The participants initially experienced difficulty following English subtitles due to their rapid pace and density. Many reported pausing frequently to keep up with the dialogue. Over time, however, most students adapted to the pacing and perceived improvement in their reading fluency. One mid-level participant stated, *"At first, I had to stop all the time to read the subtitles—it was too fast. But after a few episodes, I could follow without pausing."* Some students described using inferencing and keyword identification to compensate for comprehension gaps. As one high-level student explained, *"My comprehension wasn't perfect, but I didn't feel lost. I just focused on the main words to guess the meaning."*

Theme 2: Plot Engagement as a Viewing Motivator

All the participants emphasized that plot engagement, rather than the language of the drama, sustained their motivation to continue watching. A low-level participant remarked, *"If the story is interesting, I don't care whether it's English or Korean—I just want to know what happens next."* Despite this, several participants reported that American series with English captions were linguistically more challenging than Korean dramas, citing the frequent use of slang and idiomatic expressions as barriers. A mid-level student noted, *"American shows have so many slang words. Sometimes I can't catch the meaning even with subtitles."*

Theme 3: Comprehension of Korean Dramas with English Subtitles

All the interviewees relied on English subtitles when watching Korean dramas due to their limited proficiency in Korean. The students with lower vocabulary gains reported that comprehension was easier than expected, attributing this to cultural familiarity and the availability of accessible visual cues. As one low-level student reflected, *"I don't know Korean, but I can guess what's happening. The situations and emotions are familiar, so I can understand the subtitles better."* Similarly, a high-level participant emphasized the role of nonverbal information: *"I rely on characters' facial expressions and body language a lot. Even if I don't know a word, I can still get the meaning."* These comments suggest that multimodal input and shared cultural context facilitate inferencing and incidental vocabulary learning.

Theme 4: Engagement with English Subtitles in Mandarin Dramas

Most students reported minimal attention to English subtitles when watching Mandarin (L1) dramas, as comprehension of spoken content was effortless. However, some noted increased attention when subtitles corresponded to personally relevant or commonly used expressions. A mid-level student commented, *"Usually I ignore the subtitles, but if the line is something I say in real life, I'll look to see how it's expressed in English."* This suggests that relatable content promotes deeper engagement with English subtitles and supports vocabulary noticing through personal relevance.

Theme 5: Challenges of Standard Subtitling (L2 audio with L1 Subtitles)

All the participants agreed that standard subtitling—English audio accompanied by Chinese subtitles—had the least effect on vocabulary development. Learners found it difficult to associate Chinese meanings with spoken English forms. A high-level participant explained, *"When I watch English TV series with Chinese subtitles, my eyes just go to the Chinese. I don't really listen to the English words."* Others attributed this challenge to syntactic differences between the two languages or limited lexical knowledge. As one low-level student observed, *"The sentence order is so different, and sometimes I just can't match what I hear to the Chinese translation."* These difficulties highlight the cognitive demands of cross-linguistic processing and the constraints imposed by insufficient lexical thresholds.

The interviews revealed that learners gradually adapted to English subtitles through the use of strategies and exposure. Plot engagement and cultural familiarity enhanced motivation and comprehension, while linguistic and structural incongruence limited vocabulary development under standard subtitling conditions. Overall, the findings underscore the pedagogical value of different subtitling modalities in facilitating incidental vocabulary acquisition through binge-watching.

Conclusion and Implications

This research extended beyond bimodal, standard, and reversed subtitling to include L3 audio with L2 subtitles, aiming to explore their impact on the vocabulary gains of EFL learners. Given that many students spend a significant amount of time binge-watching dramas, these viewing habits may serve as a valuable and engaging avenue for promoting English vocabulary development. English-subtitled dramas, in particular, offer input with motivational qualities similar to those found in graded readers used in extensive reading programs.

The study revealed significant differences in vocabulary gains among eight groups across two proficiency levels, each exposed to four subtitling modes. Taking turns with each modality, the upper-intermediate groups acquired an average of 20.89 out of 28 target words, while the low-intermediate groups gained an average of 17.12 words. Among the subtitling types, L3 audio with L2 subtitles–despite students' limited knowledge of L3–produced the highest gains in vocabulary. The adjusted mean scores ranged from 6.621 to 6.773 words out of 7 for the upper-intermediate groups and from 5.691 to 6.004 for the low-intermediate groups. This was followed by bimodal subtitling (5.754–6.224 vs. 4.774–5.235), reversed subtitling (4.544–4.689 vs. 3.537–3.689), and, lastly, standard subtitling, which yielded the lowest gains (3.544–3.621 vs. 2.537–2.619).

Low-intermediate learners may initially find English-subtitled dramas challenging, and some level of comprehension difficulty is expected during the first few episodes. However, as learners adapt to the speed of the subtitles, their understanding tends to improve with continued exposure. Visual imagery also plays a crucial role, as it lowers the vocabulary threshold for processing English subtitles and facilitates the connection between L2 forms and L1 meanings, thereby compensating for limited lexical knowledge. Several strategies can help overcome early difficulties. Previewing trailers and reading plot summaries can build background knowledge to aid comprehension. Watching a drama first with L1 subtitles and then re-watching it with English subtitles is another effective approach. The familiarity gained during the first viewing may make it easier to follow the storyline in the second. Additionally, encouraging learners to re-watch English-subtitled dramas they enjoy can promote repeated encounters with unfamiliar words, thereby increasing the likelihood of vocabulary acquisition.

Admittedly, this study focused primarily on individual word learning. However, learners are also likely to attend to recurring multiword expressions while watching English-subtitled content. Although the effect of subtitling on the acquisition of such expressions lies beyond the scope of this study, it presents a promising direction for future research. As with graded readers, on-screen text does not require a radically new methodology for vocabulary acquisition. Instead, binge-watching habits may act as a catalyst for sustained engagement with English subtitles. The aim of this research has been to highlight and raise awareness of this underutilized potential.

References

Bairstow, D., & Lavaur, J. M. (2011). Audiovisual information processing by monolinguals and bilinguals: Effects of intralingual and interlingual subtitles. In A. Remael, P. Orero, & M. Carroll (Eds.), *Audiovisual translation and media accessibility at the crossroads: Media for all 3* (pp. 273–293). Rodopi.

Čepon, S. (2011). Interlingual subtitling as a mode of facilitating incidental foreign language acquisition. *English for Specific Purposes World, 11*(33), 1–36.

Danan, M. (1992). Reversed subtitling and dual coding theory: New directions for foreign language instruction. *Language Learning, 42*(4), 497–527. https://doi.org/10.1111/j.1467-1770.1992.tb01042.x

Danan, M. (2004). Captioning and subtitling: Undervalued language learning strategies. *Meta: Journal des Traducteurs, 49*(1), 67–77. https://doi.org/10.7202/009021ar

Day, R. R., & Bamford, J. (1998). *Extensive reading in the second language classroom.* Cambridge University Press.

Dizon, G. (2016). Online video streaming in the L2 classroom: Japanese students' opinions towards Netflix and subtitles. *Osaka JALT Journal, 3*, 70–87. https://static1.1.sqspcdn.com/static/f/462558/27025628/1463323273223/Osaka_JALT_J

Dizon, G. (2018). Netflix and L2 learning: A case study. *The EuroCALL Review, 26*(2), 30–40. https://doi.org/10.4995/eurocall.2018.9080

Dizon, G. (2021). Subscription video streaming for informal foreign language learning: Japanese EFL students' practices and perceptions. *TESOL Journal, 12*(2), e566. https://doi.org/10.1002/tesj.566

Fazilatfar, A. M., Ghorbani, S., & Samavarchi, L. (2011). The effect of standard and reversed subtitling versus no subtitling mode on L2 vocabulary learning. *Teaching English as a Second Language Quarterly, 30*(1), 43–64. https://doi.org/10.22099/jtls.2012.393

Gorjian, B. (2014). The effect of movie subtitling on incidental vocabulary learning among EFL learners. *International Journal of Asian Social Science, 4*(9), 1013–1026. https://www.aessweb.com/journals/5007

Grabe, W. (2009). *Reading in a second language: Moving from theory to practice.* Cambridge University Press.

Holobow, N. E., Lambert, W. E., & Sayegh, L. (1984). Pairing script and dialogue: Combinations that show promise for second or foreign language learning. *Language Learning, 34*(4), 59–74. https://doi.org/10.1111/j.1467-1770.1984.tb00352.x

Horst, M., Cobb, T., & Meara, P. (1998). Beyond a clockwork orange: Acquiring second language vocabulary through reading. *Reading in a Foreign Language, 11*(2), 207–223. https://doi.org/10.64152/10125/66953

Hsu, W. (2023a). Extensive reading through watching English-subtitled K-dramas. *THAITESOL Journal, 36*(1), 46–65. https://so05.tci-thaijo.org/index.php/thaitesoljournal/article/view/266026

Hsu, W. (2023b). Binge-watching TV series: English subtitles as an alternative to graded readers for extensive reading. *English Language Teaching Educational Journal, 6*(3), 151–161. https://doi.org/10.12928/eltej.v6i3.10068

Kim, N. (2020). Languages on the screen: A study on reversed subtitling and EFL receptive language skills. *Multimedia-Assisted Language Learning, 23*(1), 37–66. https://typeset.io/journals/multimedia-assisted-language-learning-1g5ayf2d

Krashen, S. (2004). The case for narrow reading. *Language Magazine, 3*(5), 17–19. https://www.researchgate.net/publication/265540656_The_Case_for_Narrow_Reading

Krashen, S. (2011). *Free voluntary reading.* Heinemann.

Kwarteng, A. H. (2023). Incidental vocabulary acquisition through watching movies with bimodal and standard subtitles: The case of L2 Swahili learners. *Journal for Language Teaching, 57*(2), Article 5935. https://doi.org/10.56285/jltVol57iss2a5935

Lambert, W. E., Boehler, I., & Sidoti, N. (1981). Choosing the languages of subtitles and spoken dialogues for media presentations: Implications for second language education. *Applied Psycholinguistics, 2*(2), 133–148. https://doi.org/10.1017/S0142716400000904

Locher, M. A., & Messerli, T. C. (2020). Translating the other: Communal TV watching of Korean TV drama. *Journal of Pragmatics, 170*, 20–36. https://doi.org/10.1016/j.pragma.2020.07.002

Markham, P., & Peter, L. (2003). The influence of English language and Spanish language captions on foreign language listening/reading comprehension. *Journal of Educational Technology Systems, 31*(3), 331–341. https://doi.org/10.2190/BHUH-420B-FE23-ALA0

Milton, J. (2010). The development of vocabulary breadth across the CEFR levels. In I. Bartning, M. Martin, & I. Vedder (Eds.), *Communicative proficiency and linguistic development: Intersections between SLA and language testing research* (pp. 211–232). Eurosla.

Nation, I. S. P. (2013). *Learning vocabulary in another language* (2nd ed.). Cambridge University Press.

Nation, I. S. P. (2014). How much input do you need to learn the most frequent 9,000 words? *Reading in a Foreign Language, 26*(2), 1–16. https://doi.org/10.64152/10125/66881

Nation, I. S. P. (2015). *Principles guiding vocabulary learning through extensive reading.* University of Hawaii, National Foreign Language Resource Center.

Nesirli, A. (2025). Exploring the modalities of audiovisual translation: Focus on dubbing and subtitles. *Porta Universorum, 1*(4), 6–21. https://doi.org/10.69760/portuni.0104001

Netflix. (2014, December 13). *Netflix declares binge watching is the new normal.* https://www.prnewswire.com/news-releases/netflix-declares-binge-watching-is-the-new-normal-235713431.html

Paivio, A. (1971). *Imagery and verbal processes.* Holt, Rinehart and Winston.

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

Pedersen, J. (2019). Fansubbing in subtitling land: An investigation into the nature of fansubs in Sweden. *Target, 31*(1), 50–76. https://doi.org/10.1075/target.18017.ped

Reynolds, B. L., Cui, Y., Kao, C., & Thomas, N. (2022). Vocabulary acquisition through viewing captioned and subtitled video: A scoping review and meta-analysis. *Systems, 10*(5), 133–152. https://doi.org/10.3390/systems10050133

Rubenking, B., & Bracken, C. C. (2021). Binge watching and serial viewing: Comparing new media viewing habits in 2015 and 2020. *Addictive Behaviors Reports, 14*, 100356. https://doi.org/10.1016/j.abrep.2021.100356

Talaván Zanón, N. (2006). Using subtitles to enhance foreign language learning. *Porta Linguarum, 6*, 41–52. https://doi.org/10.30827/digibug.30659

Talaván Zanón, N. (2007). Learning vocabulary through authentic video and subtitles. *TESOL-Spain Newsletter, 31*, 5–8. https://www.academia.edu/2571090/Learning_Vocabulary_through_Authentic_Video_and_Subtitles

Webb, S. (2008). The effects of context on incidental vocabulary learning. *Reading in a Foreign Language, 20*(2), 232–245. https://doi.org/10.64152/10125/66826

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), 34–70. https://doi.org/10.1075/itl.168.1.02web

Zahar, R., Cobb, T., & Spada, N. (2001). Acquiring vocabulary through reading: Effects of frequency and contextual richness. *Canadian Modern Language Review, 57*(4), 541–572. https://doi.org/10.3138/cmlr.57.4.541

About the Author

Wenhua Hsu is currently teaching at I-Shou University in Taiwan. She holds a PhD in the field of ESP from the University of Essex, UK, and an MBA from Kansas State University, USA. Her interests include frequent academic/sub-technical and lay-technical vocabulary as well as lexical bundles in specialized fields.