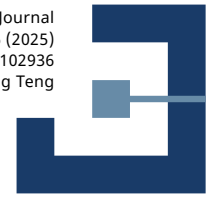
The JALT CALL Journal
ISSN 1832-4215, 21(2), 102936 (2025)
https://doi.org/10.29140/jaltcall.v21n2.102936



Castledown

Validation of Vocabulary Learning Strategies in Watching Audiovisual Materials

Mark Feng Teng
Macao Polytechnic University, Macau SAR, China
markteng@mpu.edu.mo

The present study aims to bridge the gap of validating the Vocabulary Learning Strategies in Watching Audiovisual Materials (VLSWAM) questionnaire, a tool designed to assess strategies learners use to acquire vocabulary through audiovisual content. Grounded in multimedia learning theory, the study employed exploratory factor analysis (EFA) and confirmatory factor analysis (CFA) to identify and refine the questionnaire's structure. The results revealed a robust four-factor model encompassing emotion, attention, metacognition, and information processing. These dimensions align with established frameworks, such as the Selection-Organization-Integration (SOI) model and the Trans-Symbolic Comprehension (TSC) framework, highlighting the interplay of cognitive, metacognitive, and affective strategies in multimedia learning. Practical implications include explicit training in metacognitive strategies, affective support through captions, and activities fostering multimodal integration. Particularly relevant for English as a Foreign Language (EFL) learners, the validated questionnaire contributes to the literature by offering a nuanced understanding of how learners engage with audiovisual materials for vocabulary acquisition.

Keywords: vocabulary learning, strategies, audiovisual materials, EFA, CFA

Introduction

Audiovisual materials, such as videos, have emerged as powerful instructional tools in language learning, offering a multisensory experience that combines motivational, attentional, and affective elements to engage learners. Platforms like YouTube, for instance, have revolutionized how individuals access and interact with language, providing authentic contexts where visual and auditory cues work together to facilitate comprehension (Baltova, 1994). These audiovisual materials enhance auditory processing and provide rich contextual input, closely mirroring real-life scenarios where visual cues complement spoken language, enabling viewers to "see" and "hear" messages simultaneously (Danan, 2004). Captioned videos, in particular, amplify learning by exposing students

to online texts, improving comprehension (Montero Perez, 2020) and vocabulary acquisition (Teng, 2021). The multisensory nature of audiovisual media, combining visual and auditory stimuli, boosts engagement, aids retention, and creates dynamic, immersive learning experiences tailored to learners' cognitive and affective needs (Vanderplank, 2016; Vanderplank & Teng, 2024).

While the pedagogical value of audiovisual materials is widely recognized, there remains a critical gap in understanding the specific strategies learners use to acquire vocabulary while engaging with these resources (List, 2018). Existing studies on video processing often focus narrowly on comprehension of single videos, overlooking the nuanced vocabulary learning strategies that learners employ as they watch and interact with multiple audiovisual texts (Papin & Kaplan-Rakowski, 2024). While listening enhances vocabulary learning (Vidal, 2003), audiovisual materials also offer potential (Teng, 2021) or may be the effects of audiovisual materials are more significant than listening (Teng, 2024a, 2025a). Viewing itself is not passive decoding but an active cognitive process requiring prediction and inference to compensate for incomplete auditory input (Danan, 2004). Moreover, vocabulary learning is not merely a byproduct of passive viewing or listening; it involves active cognitive processes such as prediction, inference, and the integration of contextual cues (Vanderplank & Teng, 2024). Though visual cues in videos may not directly clarify spoken language, captions are beneficial for vocabulary-related skills: they enhance verbatim recall, contextual reuse of words, and communicative performance (Neuman & Koskinen, 1992; Vanderplank, 1988). To fully harness the pedagogical potential of audiovisual materials, it is essential to examine how learners use strategies for extracting and integrating information from these resources.

The present study centers on the validation of Vocabulary Learning Strategies in Watching Audiovisual Materials (VLSWAM) questionnaire, which is designed to systematically investigate how English as a Foreign Language (EFL) learners approach vocabulary acquisition through audiovisual content. By exploring the strategies students use, such as utilizing captions, leveraging visual context, and synthesizing information across multiple videos, the present study aims to uncover effective practices that support vocabulary retention and deeper linguistic engagement. Insights gained from the VLSWAM questionnaire will inform educators and curriculum designers on how best to harness the unique strengths of audiovisual materials, ultimately optimizing vocabulary learning outcomes in diverse educational settings. By examining how EFL students process audiovisual content, particularly their strategies to enhance comprehension and vocabulary retention, researchers can identify effective practices to optimize learning outcomes. Such insights are vital for designing curricula that leverage the unique strengths of audiovisual materials in fostering deeper linguistic and cognitive engagement.

Literature review

Theoretical frameworks

The present study is grounded in a robust theoretical foundation that underscores the importance of strategic processing in multimedia learning environments. Central to this foundation is Mayer's Cognitive Theory of Multimedia Learning (CTML; Mayer, 2002), which posits that learners construct knowledge more effectively when they are able to process information from multiple modalities, such as text, audio, and visuals, in an integrated manner. According to CTML, the human cognitive system has separate channels for processing auditory/verbal and visual/pictorial information, each with limited capacity. Effective learning from multimedia, therefore, depends on learners' ability to actively manage and coordinate these channels.

Mayer's Selection-Organization-Integration (SOI) framework further elaborates on this process by delineating three essential cognitive phases. First, learners must select relevant information from the multimedia input, such as focusing on key spoken words or critical visual cues in a video. Next, they organize this information into coherent mental structures, grouping related vocabulary or concepts. Finally, they integrate new information with existing prior knowledge, for instance, by relating new vocabulary encountered in a video to previously learned words or personal experiences. These phases are not confined to a single modality but apply across auditory, visual, and combined audiovisual inputs, making them particularly relevant for understanding how learners derive meaning and acquire vocabulary from videos.

Building on CTML, Fiorella and Mayer (2013, 2015) introduced the concept of generative learning strategies, such as summarizing, self-explaining, and concept mapping, that enable learners to go beyond passive reception and actively construct deeper understanding from multimedia materials. These strategies foster meaningful connections between new information and prior knowledge, encouraging learners to process, reorganize, and elaborate on the content they encounter in audiovisual contexts. In the case of vocabulary learning, strategies might include creating mental or written summaries of new words, drawing semantic maps, or generating personal examples based on video content.

Complementing Mayer's framework, Loughlin and Alexander's (2012) Trans-Symbolic Comprehension (TSC) framework offers a nuanced perspective on the types of strategies learners use when interacting with multimedia. The TSC framework distinguishes between trans-symbolic strategies, which involve higher-order cognitive processes such as activating prior knowledge and making inferences across different symbol systems (e.g., connecting visual and verbal information), and symbol-specific strategies, which are tailored to the unique characteristics of a particular mode (e.g., interpreting visual metaphors in images or identifying keywords in captions and spoken language). In the context of video comprehension, learners may employ medium-specific tactics, such as pausing, rewinding, or replaying segments, to regulate their

cognitive load and ensure comprehension, while also drawing on more traditional text-based strategies like note-taking or keyword identification.

These theoretical frameworks illuminate the complex interplay of cognitive, metacognitive, and strategic processes involved in learning from audiovisual materials. They highlight the necessity of equipping learners with a repertoire of strategies not only to select, organize, and integrate information, but also to flexibly adapt their approach based on the affordances of different media. The VLSWAM questionnaire is thus designed to capture this multidimensional strategy use, providing a comprehensive assessment of how learners navigate, process, and internalize vocabulary in richly multimodal learning environments.

Vocabulary learning strategies

Vocabulary learning strategies encompass the wide range of techniques, approaches, and deliberate actions that learners use to facilitate the acquisition, retention, and effective use of new vocabulary items. These strategies can include activities such as guessing word meanings from context, using dictionaries, organizing vocabulary into semantic groups, employing mnemonic devices, and engaging in repeated practice or self-testing. The choice and effectiveness of these strategies can be influenced by a variety of factors, including the learner's proficiency level, learning context, motivation, and individual learning preferences.

Over the past three decades, research interest in vocabulary learning strategies has grown substantially. Early foundational studies, such as those by Gu and Johnson (1996), systematically categorized and examined the strategies used by language learners, laying the groundwork for subsequent empirical investigations. More recent work by Gu (2018) has further refined the theoretical frameworks surrounding vocabulary learning strategies and emphasized the need for rigorous validation of strategy taxonomies and measurement instruments. This growing body of research has contributed to a deeper theoretical understanding of how learners approach vocabulary learning and the specific strategies that are most effective in EFL contexts.

Research in second language acquisition highlights the importance of effective learner strategies (Teng, 2023), particularly in the use of captions and subtitles to enhance learning outcomes (Vanderplank, 2016). Contemporary perspectives characterize proficient learners as active processors of information who engage in hypothesis testing and successive approximation to construct meaning from captioned input (Teng, 2022a). Central to this process are cognitive strategies such as inferencing, contextual guessing, clarifying ambiguities, and verifying interpretations (Rubin, 1995).

Empirical evidence further underscores the significance of these strategies. For instance, Teng's (2022b) longitudinal study demonstrated that the systematic application of both cognitive and metacognitive strategies significantly improves vocabulary acquisition among young learners. Metacognitive strategies, in particular, involve the regulation of the learning process, including planning viewing sessions (e.g., managing pacing and replays), setting

comprehension goals, monitoring understanding, and adapting strategies as needed (Thompson & Rubin, 1996). Additionally, Tseng et al. (2006) investigated the construct validity of self-regulated vocabulary learning strategies through exploratory and confirmatory factor analyses. Their findings revealed that self-regulated vocabulary learning is underpinned by a general factor encompassing five subdimensions: commitment, metacognitive, satiation, emotional, and environmental strategies.

Strategies in viewing audiovisual materials

By adopting a strategic approach when interacting with audiovisual content, learners are able to maximize the language learning opportunities that such materials provide (Rubin, 1995; Thompson & Rubin, 1996). Audiovisual resources, such as films, television shows, instructional videos, and online multimedia, offer a rich and dynamic context in which multiple modes of input, spoken language, images, gestures, and written text, are integrated. This multimodal environment not only makes language more engaging and authentic but also expands the range of strategies that learners can employ to facilitate comprehension (Vanderplank, 2016) and vocabulary acquisition (Vanderplank & Teng, 2024).

Empirical studies on audiovisual materials, particularly captioning, reveal its capacity to foster active engagement and metacognitive strategy use. Vanderplank (1988) observed that learners initially found captions distracting but gradually developed adaptive strategies, such as toggling between audio and text or processing multiple channels simultaneously. In a later study, learners with intermediate-to-advanced English proficiency reported using captions as a scaffold for comprehension or a tool to self-assess listening skills (Vanderplank, 1990). However, Magliano et al. (2013) caution that strategic processing may vary across mediums, with some strategies being modality-specific. For example, Lee and List (2019) found that undergraduates relied more on selection strategies (identifying key information) during video viewing than on higher-order organizational tactics common in text processing. Other studies categorize video-specific strategies into content-based note-taking, playback control (pausing/rewinding), and navigational adjustments (Caspi et al., 2005), as well as cognitive (inferencing), memory (encoding/retrieval), and compensatory (guessing) approaches (Lin, 2011). List's (2018) comparison of video and text processing further highlights medium-dependent differences: while text prompts vocabulary-focused strategies (e.g., contextual analysis), video encourages multimodal integration. These findings align with the TSC framework, illustrating how learners blend universal and modality-specific strategies to navigate diverse symbolic systems.

Affective factors also play a pivotal role in audiovisual learning. Krashen's (1985) "affective filter" hypothesis posits that low-anxiety environments enhance information intake, a principle supported by studies on captions. Borras and Lafayette (1994) found that captioned content significantly improved learners' attitudes by reducing anxiety about missing critical

information. This reassurance allows learners to focus cognitive resources on comprehension rather than frustration (Teng, 2021). Captions also provide instant feedback, reinforcing motivation and confidence (Vanderplank, 2016). Hsu's (2015) experiment demonstrated that adaptive bilingual captions boosted intrinsic motivation and enjoyment, underscoring their role in creating engaging, learner-centered environments.

Collectively, this line of research illuminates how strategic and affective factors synergize to enhance vocabulary learning through captioned media. Proficient subtitle users, particularly in regions with widespread subtitling, often subconsciously deploy effective strategies for incidental learning (Vanderplank, 1988). However, learners unfamiliar with captions require explicit guidance to adopt intentional strategies. Educators must systematically introduce tactics that encourage reflective engagement with both auditory and written inputs, moving beyond passive listening. Assessing learners' strategy use during audiovisual tasks is critical, as multimedia environments simulate real-world communication and foster skills for vocabulary acquisition (Vanderplank & Teng, 2024). By integrating these insights, educators can empower learners to harness audiovisual materials, both in and beyond the classroom, as dynamic tools for vocabulary development.

The present study, by validating VLSWAM, contributes to the literature on vocabulary learning and multimedia instruction, as well as provides practical insights for educators seeking to harness the full potential of audiovisual materials in language education. The validated VLSWAM questionnaire will serve as a valuable tool for future research and pedagogical innovation, supporting the design of targeted interventions that foster effective vocabulary acquisition in increasingly digital and multimodal learning environments.

Rationale for the present study

Despite the growing prevalence of audiovisual materials in language learning environments, there remains a significant gap in our understanding of the specific strategies learners employ to acquire vocabulary while engaging with these resources. While previous research has established the general benefits of videos and captioned content for vocabulary development and comprehension (Montero Perez, 2020; Teng, 2024b, 2025b), researchers have not paid attention to the underlying processes and strategies learners use for vocabulary learning during audiovisual engagement. Furthermore, existing research has often examined these strategies in isolation or within limited contexts, neglecting the complex, multifaceted nature of vocabulary learning in multimedia environments (Papin & Kaplan-Rakowski, 2024).

This gap is particularly pronounced in the context of EFL learning, where students, even though with limited proficiency level, are increasingly exposed to diverse audiovisual materials both inside and outside the classroom. Although researchers have highlighted the importance of active cognitive engagement, such as prediction, inference, and synthesis of information from multimodal sources (Danan, 2004), there is a lack of comprehensive, empirically validated

instruments designed to capture the range of vocabulary learning strategies employed by EFL learners in these settings. As a result, educators and curriculum designers have limited guidance on how to effectively support and scaffold vocabulary development through audiovisual materials.

The primary objective of the present study is to address this critical gap by validating the VLSWAM Questionnaire. The VLSWAM was developed to systematically assess the multifaceted strategies learners use when encountering new vocabulary through audiovisual content, drawing on the theoretical foundation of multimedia learning theory. This framework posits that meaningful learning occurs when learners actively integrate verbal and visual information, making the investigation of strategy use particularly relevant in the context of video-based learning. To ensure the VLSWAM's robustness and theoretical alignment, the present study aims to examine its factorial structure using a rigorous two-stage approach. First, exploratory factor analyses (EFA) will be conducted to identify the underlying dimensions of vocabulary learning strategies represented in the questionnaire. This will be followed by confirmatory factor analyses (CFA) using structural equation modeling (SEM) to test the stability and validity of the identified factor structure. Through this process, the present study seeks to establish the VLSWAM as a reliable and valid instrument capable of capturing the diverse strategies learners employ when watching audiovisual materials.

The study was guided by a central research question, which drove the investigation:

Which structural model best represents the dimensions of the VLSWAM questionnaire?

Methods

Participants

A total of 606 undergraduate students participated in the study, recruited from six universities located in the southern region of China, mostly from Guangdong and Guangxi provinces. The participants were selected through convenience sampling, as the author has connection with those universities. The participants represented a diverse range of academic majors: Economics (15%, n = 110), Engineering (16%, n = 86), Visual Arts (17%, n = 85), Psychology (14%, n = 83), English (25%, n = 84), and Marketing (13%, n = 158). The age of participants ranged from 18 to 22 years, with a mean age of 20.03 years (SD = 1.19). Regarding gender distribution, 44% (n = 266) of the students were male and 56% (n = 340) were female. On average, participants reported having 9 years of formal English language learning experience (M = 9.12, SD = 1.03). Furthermore, all students indicated prior experience with watching audiovisual materials, although none had studied in an English-speaking country.

Questionnaire development

The development of the VLSWAM questionnaire involved a comprehensive three-phase process, which encompassed item generation, initial piloting, and psychometric evaluation. The first phrase was the item generation stage. To ensure the credibility and quality of the questionnaire items, the item generation phase commenced with focus group interviews, following Dörnyei's (2010) recommendation of involving targeted learners in this process. A total of 16 undergraduate students, selected to represent diversity in terms of gender, year level, and disciplinary major, were invited to participate in the focus group interviews. The purpose of these interviews was to gain insights directly from students about the strategies they used to learn vocabulary from audiovisual materials and acquire vocabulary knowledge from audiovisual input, both in and out of the classroom. During the focus group interviews, participants were encouraged to share their experiences and describe the specific strategies they utilized. Their inputs and discussions provided valuable information that guided the generation of initial items for the VLSWAM questionnaire. By incorporating the perspectives of learners who have direct experience with vocabulary learning in the context of audiovisual materials, we aimed to enhance the questionnaire's relevance and validity. The analysis of the focus group data, including the participants' descriptions of their vocabulary learning strategies, served as a foundation for generating the initial items of the VLSWAM questionnaire. This iterative process ensured that the questionnaire items captured the diverse range of strategies employed by learners when engaging with audiovisual materials for vocabulary acquisition.

The second phase in the development of the VLSWAM questionnaire involved an extensive review of the relevant literature, with a particular focus on examining references pertaining to strategies in multimedia learning. To ensure a comprehensive understanding of the topic, two researchers in the field of applied linguistics were invited to check scholarly works (Fiorella & Mayer, 2015; Magliano et al., 2013; Mayer, 2002; Vanderplank, 2016). By critically examining and synthesizing the insights from these literature sources, we aimed to incorporate established knowledge and theoretical frameworks into the item-generation process of the questionnaire, aligning with the recommendations of Dörnyei (2010), who argue that consulting existing theoretical frameworks enhances the construct validity of a questionnaire. The item generation process for the VLSWAM questionnaire resulted in a comprehensive list of 60 items specifically related to vocabulary learning strategies when watching audiovisual materials.

The final phase involved a thorough psychometric evaluation to ensure the quality and validity of the questionnaire. To commence the evaluation, three researchers with expertise in the field of vocabulary learning were invited to examine the initial list of items. Their role was to carefully scrutinize the theoretical rationale behind each item, assess whether the questions accurately measured the intended constructs, and rate the extent to which the survey questions aligned with the constructs as defined in the study. For alignment and relevance, the primary unit of measurement was a 4-point relevance scale,

where 1 indicated "not relevant," 2 "somewhat relevant," 3 "quite relevant," and 4 "highly relevant." Each researcher independently rated the extent to which each item corresponded to the construct definitions established in the study. Additionally, researchers considered the clarity, specificity, and theoretical fit of each item, evaluating whether the item's value, its contribution to construct measurement, was sufficient for inclusion. Based on their evaluation, five items that received the ratings of "not relevant" and "somewhat relevant" were eliminated from the initial list. This expert review process aimed to enhance the questionnaire's content validity and ensure the items effectively captured the desired constructs. Following this, the revised list of items was presented to a group of 10 EFL students, who were not included at a later stage for data collection. Their task was to evaluate the clarity and readability of the items. Through this process, double-barreled items were identified and removed, while related statements were consolidated or rephrased to enhance clarity. The feedback provided by the EFL students played a crucial role in refining the questionnaire and ensuring its comprehensibility for the target population. As a result of these iterative revisions and evaluations, the final version of the VLSWAM questionnaire was developed, containing a total of 49 items. Appendix A shows the final validated VLSWAM. This rigorous process of expert evaluation and student feedback aimed to enhance the validity, relevance, and overall quality of the instrument.

Procedures

Considering the participants' varying levels of English language proficiency, the original English questionnaire was translated into Chinese to ensure full comprehension and accurate responses. The translation process adhered to established guidelines, including forward and backward translation, to maintain the integrity and equivalence of the questionnaire items across both languages. To assess the trait features of the vocabulary learning strategies, a 7-point Likert scale was employed, ranging from 1 (Strongly disagree) to 7 (Strongly agree). Participants were asked to rate each item on this scale, and the mean scores of the items were cumulatively calculated to provide an overall measure of the strategies employed.

The questionnaire was administered online. Participants received a QR code, which, when scanned, provided direct access to the Chinese version of the survey. Prior to completing the questionnaire, participants were given clear written and verbal instructions. Researchers were present to clarify any instructions or answer participants' questions, ensuring that all items were understood correctly.

On average, participants took approximately 10 to 15 minutes to complete the Chinese version of the questionnaire. This duration allowed sufficient time for thoughtful consideration and accurate responses to each item. All responses were collected electronically and securely stored for subsequent analysis.

Data analysis

To determine the most suitable factor structure for representing the VLSWAM questionnaire, both exploratory factor analysis (EFA) and confirmatory factor analysis (CFA) were conducted. Prior to analysis, the raw dataset of 606 survey responses underwent a series of data management steps to ensure accuracy and usability. First, the data were carefully screened for incomplete responses, inconsistencies, and outliers. We did not find any missing data, nor outliers. Next, the responses were formatted for analysis, with categorical and Likert-scale items coded appropriately and checked for input errors. After cleaning and formatting, the dataset was randomly divided into two equal halves to facilitate independent analyses, with the first half (n = 303) used for EFA and the second half (n = 303) reserved for CFA (Bollen, 1989).

In the present study, EFA was employed to examine the underlying factor structure of the VLSWAM by exploring the relationships among the questionnaire items. Principal axis factoring with oblimin rotation was chosen as the analytical method. This decision was based on the assumption that there existed an underlying theoretical structure, and it was anticipated that the dimensions or factors describing the structure might be intercorrelated. For factor extraction, the determination of the number of factors to retain was initially guided by eigenvalues, with any factor having an eigenvalue of 1.0 or higher being retained. Additionally, a scree plot was examined to identify any changes in the slope of the line connecting the eigenvalues of the factors. Subsequently, the factors were rotated. In this study, an oblique rotation (oblimin) was employed because it was assumed that the factors would be related (Lorenzo-Seva et al., 2009). Finally, an assessment was conducted to examine if the items that loaded strongly on each factor were conceptually coherent.

Confirmatory factor analysis was then employed to validate the factor structure obtained from the EFA using the second half of the sample. Maximum likelihood estimation was used for all tested models. Model goodness of fit (GOF) was evaluated using several indices, including the Chi-square statistic, Chi-square/df ratio, Goodness-of-Fit Index (GFI), Adjusted Goodness-of-Fit Index (AGFI), Comparative Fit Index (CFI), Root-Mean-Square Error of Approximation (RMSEA), and Tucker-Lewis Index (TLI). GFI, AGFI, TLI, and CFI values typically range from 0 to 1.0, with values of 0.90 or higher indicating a good fit (Schumacker & Lomax, 1996). An RMSEA value below 0.06 is indicative of a good model fit (Hu & Bentler, 1999). When models are fully nested, the Chi-square difference test can be employed, comparing the Chi-square values of the two models with the difference evaluated as a Chi-square statistic using the degrees of freedom difference between the two models (Bollen, 1989). Data analyses were through R (version r4.4.1).

Results

Reliability analysis

The reliability of the Vocabulary Learning Strategies in Watching Audiovisual Materials (VLSWAM) Questionnaire was assessed using Cronbach's alpha. The internal consistency of the questionnaire was excellent, with a raw alpha of 0.98 and a standardized alpha of 0.98. The mean inter-item correlation was 0.47, and the signal-to-noise ratio (S/N) was 43, indicating high reliability. The G6(smc) value (0.98) further confirmed strong item homogeneity. Confidence intervals for Cronbach's alpha (Feldt: 0.97–0.98; Duhachek: 0.97–0.98) demonstrated stable reliability estimates. The mean score across all items was 4.5 (*SD* = 0.64), with a median inter-item correlation of 0.47.

Exploratory Factor Analysis (EFA)

The dataset was split into two equal halves (303 observations each) for exploratory and confirmatory factor analyses. EFA was conducted on the first half using principal axis factoring (PAF) with oblimin rotation, assuming correlated

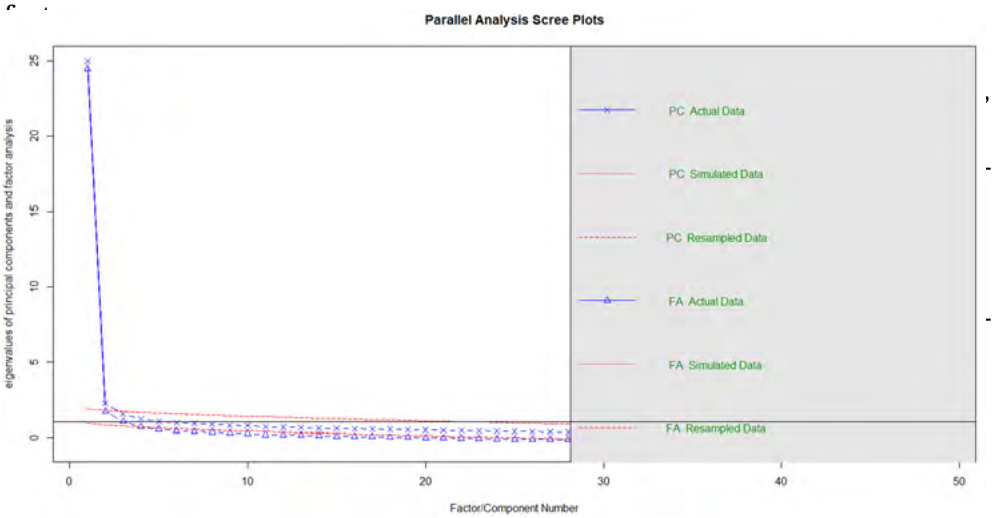


Figure 1. Parallel analysis scree plots

Table 1. Variance explained by extracted factors

Factor	SS Loadings	Proportion Var	Cumulative Var
PA3	7.06	0.14	0.14
PA1	6.74	0.14	0.28
PA4	6.53	0.13	0.41
PA2	5.89	0.12	0.53
PA5	2.80	0.06	0.59

Table 2. Selected factor loadings (pattern matrix)

Item	PA3	PA1	PA4	PA2	PA5	h^2
Q4	0.06	0.04	0.09	0.75	−0.01	0.74
Q5	0.17	−0.08	0.05	0.74	−0.01	0.70
Q15	0.03	0.72	0.14	−0.06	0.03	0.67
Q39	0.67	0.10	0.19	−0.09	0.05	0.70
Q45	0.00	0.03	0.79	−0.01	0.01	0.66

Note. Bold values indicate primary factor loadings.

The final step was to understand factor loadings and structure. Five factors were extracted and rotated obliquely (see Table 2). Loadings ≥ 0.40 were considered significant. Factor correlations ranged from 0.37 to 0.64, confirming moderate to strong inter-factor relationships.

- Factor 1 (PA3) explained 14% of the variance (eigenvalue = 7.06) and included items with high loadings related to *emotion* (e.g., q28: 0.54, q39: 0.67, q40: 0.67).
- Factor 2 (PA1) accounted for 14% of the variance (eigenvalue = 6.74) and reflected *attention* (e.g., q15: 0.72, q17: 0.55, q20: 0.50).
- Factor 3 (PA4) explained 13% of the variance (eigenvalue = 6.53) and represented *metacognition* (e.g., q45: 0.79, q46: 0.77, q43: 0.60).
- Factor 4 (PA2) contributed 12% of the variance (eigenvalue = 5.89) and captured information processing (e.g., q4: 0.75, q5: 0.74, q2: 0.64).
- Factor 5 (PA5) explained 6% of the variance (eigenvalue = 2.80) and included items related to *memory* (e.g., q10: 0.34, q13: 0.39, q33: 0.33).

Communalities (h^2) ranged from 0.42 to 0.74, indicating moderate to strong shared variance across items. The mean item complexity was 2.0, reflecting moderate multidimensionality.

We also checked the model fit for EFA. The root mean square of residuals (RMSR) was 0.03, and the df-adjusted RMSR was 0.03, indicating excellent model fit (TLI = 0.900, CFI = 0.912, $\chi 2 = 2122$, $df = 986$, $\chi 2/df = 2.152$, $p < .001$). Factor score adequacy was strong, with correlations between regression-based scores and factors ranging from 0.88 (PA5) to 0.96 (PA3, PA4). Overall, the EFA results supported a 5-factor structure for the VLSWAM questionnaire (see Table 3), with robust reliability and factor score validity. Further validation using CFA on the second half of the data is recommended.

Table 3. Item communality and final exploYour image of the ratory factor analysis results

	Factor loadings					Communality		Complexity
						Part that can be jointly explained	Part that cannot be jointly explained	Used to measure the complexity of each variable in the factor structure
	Emotion	Attention	Metacognition	Information processing	Memory	h2	u2	com
Item1				0.37		0.46	0.54	2.7
Item2				0.64		0.56	0.44	1.3
Item3				0.52		0.56	0.44	2.1
Item4				0.75		0.74	0.26	1.1
Item5				0.74		0.7	0.3	1.1
Item6				0.45	0.31	0.61	0.39	2.1
Item7	0.31			0.52		0.7	0.3	2.1
Item8			0.35			0.49	0.51	3
Item9				0.39	0.31	0.61	0.39	2.6
Item10					0.34	0.54	0.46	3.2
Item11			0.41			0.44	0.56	2.3
Item12		0.34	0.38			0.42	0.58	2.7
Item13					0.39	0.53	0.47	2.3
Item14		0.41	0.4			0.59	0.41	2.1
Item15		0.72				0.67	0.33	1.1
Item16		0.41		0.41		0.58	0.42	2.3
Item17		0.55				0.61	0.39	1.5
Item18		0.52				0.53	0.47	1.4
Item19		0.46				0.5	0.5	1.6
Item20		0.5				0.6	0.4	1.5
Item21			0.39			0.59	0.41	2.6
Item 22	0.46					0.59	0.41	2
Item 23		0.49				0.58	0.42	1.7
Item 24		0.47				0.5	0.5	1.7
Item 25						0.6	0.4	3.8
Item 26						0.56	0.44	3.4
Item 27	0.34					0.42	0.58	2.3
Item 28	0.54					0.59	0.41	1.3
Item 29		0.44				0.56	0.44	2.3
Item 30			0.35			0.6	0.4	2.9
Item 31	0.48					0.63	0.37	1.9
Item 32	0.43					0.67	0.33	2.4
Item 33		0.43			0.33	0.58	0.42	2.8
Item 34		0.43			0.36	0.63	0.37	2.3
Item 35		0.4				0.63	0.37	2.4
Item 36	0.56					0.69	0.31	1.6
Item 37	0.47					0.65	0.35	2.1
Item 38	0.54					0.65	0.35	1.9
Item 39	0.67					0.7	0.3	1.3
Item 40	0.67					0.68	0.32	1.1
Item 41	0.5		0.32			0.6	0.4	1.9
Item 42	0.51					0.58	0.42	1.5
Item 43			0.6			0.57	0.43	1.4
Item 44			0.67			0.62	0.38	1.3
Item 45			0.79			0.66	0.34	1
Item 46			0.77			0.72	0.28	1.1
Item 47	0.35		0.4			0.55	0.45	2.6
Item 48			0.49			0.64	0.36	1.6
Item 49			0.36			0.55	0.45	2.6

Confirmatory Factor Analysis (CFA)

The study employed CFA to determine whether the factor structure identified through EFA could be confirmed in a second half-sample. Structural equation modeling (SEM) methods (Arbuckle, 1997) were used to estimate the factor models.

The 5-factor model, derived from EFA, could not be validated using CFA because it failed to converge. This means that the statistical process used to estimate the model's parameters could not find a solution, indicating that the 5-factor structure does not adequately fit the data or is too complex to be reliably estimated. As a result, alternative models were tested: a model based on the first four factors identified in the EFA was tested. The model showed an improvement in fit compared to the 5-factor model, but it did not achieve optimal fit based on the goodness-of-fit indices. While some indices were within acceptable ranges, the overall fit was not ideal, indicating that the 4-factor structure from the EFA required further refinement. A model based on the first three factors identified in the EFA was also tested. This model demonstrated a better fit than the 5-factor model and the 4-factor model based on the raw EFA results. However, despite the improvement, the fit was still not optimal, suggesting that further adjustments or refinements to the factor structure were necessary. We finally developed a refined 4-factor model based on the EFA results and further modifications. Items with high complexity (i.e., those loading onto multiple factors) were removed to improve model fit. Items that could be explained by two or three dimensions were excluded, as such items reduce the predictive ability of the model. The refined model showed the best fit among all tested models. Table 4 summarizes the fit indices for the tested models.

Table 4. Model comparison results

Model	χ2χ2	*df*	*p*	GFI	AGFI	CFI	RMSEA	TLI	Δχ2Δχ2	Δ*Δdf*	*p*
5-factor model	2942.174	1127	0	0.651	0.621	0.804	0.073	0.796	-	-	-
3-factor model	1582.547	658	0	0.767	0.737	0.866	0.068	0.857	1359.627	469	0
4-factor model	2082.903	935	0	0.754	0.728	0.864	0.064	0.856	859.271	192	0
Refined 4-factor model	1193.306	590	0	0.917	0.893	0.905	0.058	0.909	1748.868	537	0

Table 4 showed the following results:

- Initial 1-Factor Model (f1): Poor fit with low GFI (0.651), AGFI (0.621), and CFI (0.804). RMSEA (0.073) exceeded the threshold for good fit.
- 3-Factor Model (f3): Improved fit compared to f1, with higher GFI (0.767), AGFI (0.737), and CFI (0.866). RMSEA (0.068) was closer to acceptable levels.

- 4-Factor Model (f4): Further improvement in fit indices compared to f3, with GFI (0.754), AGFI (0.728), and CFI (0.864). RMSEA (0.064) met the threshold for good fit.
- Refined 4-Factor Model (f4g): The best-fitting model, with GFI (0.917), AGFI (0.893), and CFI (0.905) exceeding the thresholds for good fit. RMSEA (0.058) indicated a strong fit, and TLI (0.909) was satisfactory.

The data did not support the 5-factor model. One reason could be low item loadings for the memory factor (see Table 3). The 3-factor model was also close to acceptable levels. Compared to the 3-factor or 5-factor models, the refined 4-factor model (see Figure 2) demonstrated the best overall fit among all tested models. It satisfied the majority of the goodness-of-fit criteria and outperformed the other models in terms of GFI, AGFI, CFI, and RMSEA. This model, encompassing emotion, attention, metacognition, and information processing, is considered the most optimal and valid representation of the factor structure (see Appendix A for the final survey).

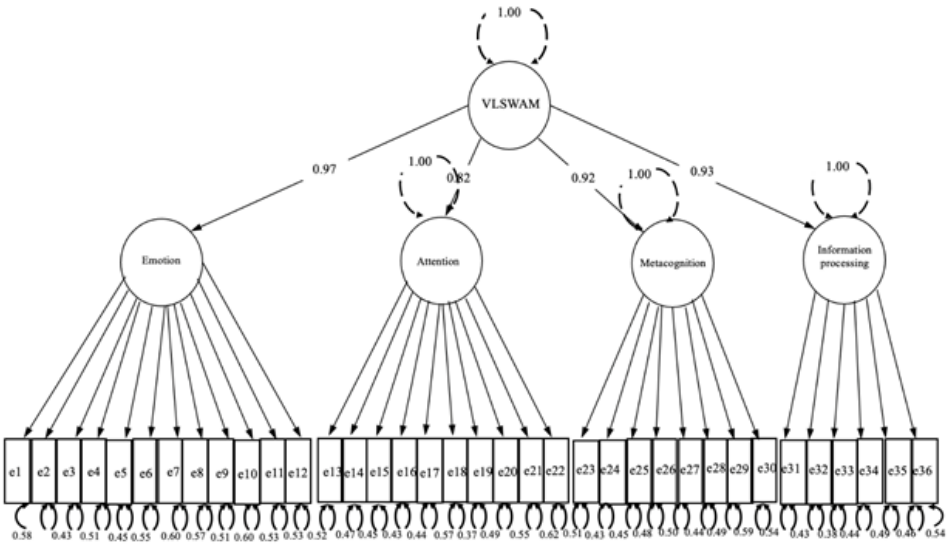


Figure 2. The refined 4-factor model

Discussion

The present study sought to validate the multifaceted structure of the Vocabulary Learning Strategies in Watching Audiovisual Materials (VLSWAM) Questionnaire, grounded in multimedia learning theory (Mayer, 2002). The findings revealed critical insights into the strategic dimensions learners employ when engaging with audiovisual content, while also highlighting complexities in modeling these strategies. Below, we contextualize these results within existing literature, discuss their theoretical and practical implications, and address limitations and future directions.

The EFA initially supported a 5-factor structure (emotion, attention,

metacognition, information processing, and memory), aligning with Mayer's (2002) Selection-Organization-Integration (SOI) framework and the Trans-Symbolic Comprehension (TSC) model (Loughlin & Alexander, 2012). The validated models posit that effective multimedia learning involves the dynamic interplay of cognitive, metacognitive, and affective strategies, enabling learners to select relevant information, organize it coherently, and integrate it with prior knowledge (Fiorella & Mayer, 2015; Mayer, 2002). The identified factors reflect this multidimensional engagement, supporting the view that successful navigation of multimodal inputs requires learners to deploy a range of strategies (Papin & Kaplan-Rakowski, 2024).

The CFA further provided evidence for the validation process. The initial 5-factor model failed to converge, necessitating refinement. The optimal solution—a four-factor model excluding memory—suggests that memory-related strategies (e.g., encoding, retrieval) may overlap with information processing or metacognitive dimensions in practice. This aligns with Lin's (2011) categorization of memory strategies as embedded within broader cognitive processes. The refined model's strong fit (GFI = 0.917, RMSEA = 0.058) underscores the interrelated nature of these constructs, where emotion, attention, metacognition, and information processing synergize to mediate vocabulary acquisition in multimedia environments. The prominence of metacognition (e.g., planning viewing sessions, self-monitoring) and information processing (e.g., contextual guessing, integrating visual cues) resonates with Fiorella and Mayer's (2015) emphasis on generative strategies, which facilitate deeper processing and transfer of learning, and with Vanderplank's (2016) observations of adaptive caption use as a form of strategic engagement. The metacognition and information processing factors reflect learners' ability to regulate their learning and actively engage with content. Metacognitive strategies, such as planning, monitoring, and evaluating comprehension, play a pivotal role in optimizing learning outcomes (Thompson & Rubin, 1996). Information processing strategies, such as inferencing and contextual guessing, further support learners in decoding and integrating new vocabulary, aligning with research highlighting the importance of active processing in multimedia learning (Fiorella & Mayer, 2015; Papin & Kaplan-Rakowski, 2024; Teng, 2021), which highlights the importance of strategic engagement with multimodal inputs for vocabulary learning. The identification of emotion as a distinct factor underscores the affective dimensions highlighted by Krashen (1985), where reduced anxiety and enhanced motivation facilitate deeper engagement with audiovisual content. The attention factor captures the cognitive effort learners invest in focusing on and processing audiovisual materials, a construct supported by studies emphasizing the role of attentional control in multimedia learning environments (e.g., Hsu, 2015; Vanderplank, 2016; Vanderplank & Teng, 2024).

Overall, these findings are consistent with a growing body of literature on multimedia learning strategies (Fiorella & Mayer, 2015; Papin & Kaplan-Rakowski, 2024; Teng, 2021; Vanderplank & Teng, 2024), which underscores the importance of strategic, affective, and attentional engagement with multimodal inputs for effective vocabulary learning and overall language development.

The validated four-factor structure of the VLSWAM questionnaire makes a significant theoretical contribution to the growing body of literature on vocabulary learning strategies in multimedia contexts. By integrating insights from the Cognitive Theory of Multimedia Learning and the Trans-Symbolic Comprehension (TSC) framework, the present study offers a nuanced understanding of how learners navigate the multimodal demands of audiovisual materials. The identification of emotion, attention, metacognition, and information processing as core dimensions not only advances theoretical perspectives on the strategic and affective processes involved in vocabulary acquisition but also provides educators with a robust tool to assess and foster learners' strategic engagement with audiovisual materials. For instance, explicit metacognitive training, such as teaching learners to plan their viewing sessions (e.g., setting goals) and regulate their learning (e.g., pausing to review captions), could significantly enhance their ability to extract linguistic input from videos. Additionally, providing affective support by integrating bilingual captions or reducing playback speed could help lower learners' anxiety (Borras & Lafayette, 1994), thereby freeing cognitive resources for comprehension. Furthermore, designing activities that encourage multimodal integration, such as analyzing video scenes with captions to synthesize visual, auditory, and textual cues, aligns with the information processing strategies identified in the questionnaire. These strategies are particularly relevant in EFL contexts, where learners often lack immersive language environments.

Limitations, future directions, and conclusion

While this study advances understanding of audiovisual learning strategies, several limitations warrant consideration. First, the sample comprised Chinese undergraduates, potentially limiting generalizability to other cultural or age groups. Second, self-reported data may introduce response bias; observational or longitudinal studies could complement questionnaire findings by capturing learners' strategy use. Third, the exclusion of memory as a standalone factor invites further exploration. Qualitative methods might elucidate how learners conceptualize memory within multimedia contexts. Future research should also investigate the causal relationships between strategy use and vocabulary outcomes. For example, experimental designs could test whether interventions targeting metacognition or emotion directly improve retention. Additionally, expanding the VLSWAM framework to include collaborative strategies (e.g., peer discussions of video content) might reflect the social dimensions of multimedia learning.

This study underscores the multifaceted nature of vocabulary learning strategies in audiovisual contexts, bridging theoretical frameworks like the SOI model with practical applications for EFL education. The refined four-factor VLSWAM questionnaire offers a validated instrument to assess learners' strategic engagement, informing pedagogical practices that harness the unique affordances of multimedia materials. By prioritizing metacognitive, affective,

and integrative strategies, educators can cultivate learners' autonomy and proficiency in vocabulary learning in an increasingly digital world.

References

Baltova, I. (1994). The impact of video on the comprehension skills of core French students. *The Canadian Modern Language Review, 50*(3), 507–532. https://doi.org/10.3138/cmlr.50.3.50

Bollen, K. A. (1989). Structural equations with latent variables. John Wiley.

Borras, I., & Lafayette, R. C. (1994). Effects of multimedia courseware subtitling on the speaking performance of college students of French. *Modern Language Journal, 78,* 61–75. https://doi.org/10.1111/j.1540-4781.1994.tb02015.x

Caspi, A., Gorsky, P., & Privman, M. (2005). Viewing comprehension: Students' learning preferences and strategies when studying from video. *Instructional Science, 33,* 31–47. https://doi.org/10.1007/s11251-004-2576-x

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

Dörnyei, Z. (2010). *Questionnaires in second language research: Construction, administration, and processing.* Routledge.

Fiorella, L., & Mayer, R. E. (2013). The relative benefits of learning by teaching and teaching expectancy. *Contemporary Educational Psychology, 38*(4), 281–288. https://doi.org/10.1016/j.cedpsych.2013.06.001

Fiorella, L., & Mayer, R. E. (2015). *Learning as a generative activity.* Cambridge University Press.

Gu, Y. (2018). Validation of an online questionnaire of vocabulary learning strategies for ESL learners. *Studies in Second Language Learning and Teaching, 8*(2), 325–350. https://doi.org/10.14746/ssllt.2018.8.2.7

Gu, Y., & Johnson. R. K. (1996). Vocabulary learning strategies and language learning outcomes. *Language Learning, 46*(4), 643–679. https://doi.org/10.1111/j.1467-1770.1996.tb01355.x

Hsu, C. K. (2015). Learning motivation and adaptive video caption filtering for EFL learners using handheld devices. *ReCALL, 27*(1), 84–103. https://doi.org/10.1017/S0958344014000214

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling, 6,* 1–55. https://doi.org/10.1080/10705519909540118

Krashen, S. (1985). *The input hypothesis: Issues and implications.* Longman.

Lee, H. Y., & List, A. (2019). Processing of texts and videos: A strategy-focused analysis. *Journal of Computer Assisted Learning, 35*(2), 268–282. https://doi.org/10.1111/jcal.12328

Lin, L. F. (2011). Gender differences in L2 comprehension and vocabulary learning in the video-based CALL program. *Journal of Language Teaching and Research, 2*(2), 295–301. https://doi.org/10.4304/jltr.2.2.295-301

List, A. (2018). Strategies for comprehending and integrating texts and vídeos. *Learning and Instruction, 57,* 34–46. https://doi.org/10.1016/j.learninstruc.2018.01.008

Lorenzo□Seva, U., van de Velden, M., & Kiers, H. A. (2009). Oblique rotation in correspondence analysis: A step forward in the search for the simplest interpretation. *British Journal of Mathematical and Statistical Psychology, 62*(3), 583–600. https://doi.org/10.1348/000711008X368295

Loughlin, S. M., & Alexander, P. A. (2012). Explicating and exemplifying empiricist and cognitivist paradigms in the study of human learning. In L. L'Abate (Ed.), *Paradigms in theory construction* (pp. 273–296). Springer.

Magliano, J. P., Loschky, L. C., Clinton, J. A., & Larson, A. M. (2013). Is reading the same as viewing? An exploration of the similarities and differences between processing text- and visually-based narratives. In B. Miller, L. Cutting, & P. McCardle (Eds.), *Unraveling the behavioral, neurobiological, and genetic components of reading comprehension* (pp. 78–90). Brookes Publishing Co.

Mayer, R. E. (2002). Multimedia learning. *Psychology of Learning and Motivation, 41*(2), 85–139. https://doi.org/10.1016/S0079-7421(02)80005-6

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

Neuman, S. B., & Koskinen, P. (1992). Captioned television as 'comprehensible input': Effects of incidental word learning from context for language minority students. *Reading Research Quarterly, 27,* 95–106. https://doi.org/10.2307/747835

Papin, K., & Kaplan-Rakowski, R. (2024). A study of vocabulary learning using annotated 360 pictures. *Computer Assisted Language Learning, 37*(5–6), 1108–1135. https://doi.org/10.1080/09588221.2022.2068613

Rubin, J. (1995): Learner processes and learner strategies. In V. Galloway & C. Herron (Eds.), *Research within research II* (pp. 129-148). Colson Printing Company.

Schumacker, R. E., & Lomax, R. G. (1996). *A beginner's guide to structural equation modeling.* Lawrence Erlbaum.

Teng, M. F. (2021). *Language learning through captioned videos: Incidental vocabulary acquisition*. Routledge.

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

Teng, M. F. (2022b). Exploring awareness of metacognitive knowledge and acquisition of vocabulary knowledge in primary grades: A latent growth curve modelling approach. *Language Awareness, 31*(4), 470–494. https://doi.org/10.1080/09658416.2021.1972116

Teng, M. F. (2023). Language learning strategies. In Z. Wen, B. Adriana, B., Mailce, & F. Teng (Eds.), *Cognitive individual differences in second language acquisition: Theory, assessment & pedagogy* (pp. 147–173). De Gruyter Mouton.

Teng, M. F. (2024a). Working memory and prior vocabulary knowledge in incidental vocabulary learning from listening, reading, reading-while-listening, and viewing captioned videos. *System, 124*, Article 103381. https://doi.org/10.1016/j.system.2024.103381

Teng, M. F. (2024b). Incidental vocabulary learning from captioned video genres: Proficiency, working memory, and aptitude. *Computer Assisted Language Learning,* 1–43. https://doi.org/10.1080/09588221.2024.2421517

Teng, M. F. (2025a). Incidental vocabulary learning from listening, reading, and viewing captioned videos: frequency and prior vocabulary knowledge. *Applied Linguistics Review, 16*(1), 477–507. https://doi.org/10.1515/applirev-2023-0106

Teng, M. F. (2025b). Effectiveness of captioned videos for incidental vocabulary learning and retention: The role of working memory. *Computer Assisted Language Learning, 38*(1–2), 206–234. https://doi.org/10.1080/09588221.2023.2173613

Tseng, W., Dörnyei, Z., & Schmitt, N. (2006). A new approach to assessing strategic learning: The case of self-regulation in vocabulary acquisition. *Applied Linguistics, 27,* 78–102. https://doi.org/10.1093/applin/ ami046

Thompson, I., & Rubin, J. (1996). Can strategy instruction improve listening comprehension? *Foreign Language Annals, 29*(3), 331–342. https://doi.org/10.1111/j.1944-9720.1996.tb01246.x

Vanderplank, R. (1988). The value of teletext subtitles in language learning. *ELT Journal, 42*(4), 272–281. https://doi.org/10.1093/elt/42.4.272

Vanderplank, R. (1990). Paying attention to the words: Practical and theoretical problems in watching television programmes with uni-lingual (CEEFAX) sub-titles. *System, 18*(2), 221–234. https://doi.org/10.1016/0346-251X(90)90056-B

Vanderplank, R. (2016). *Captioned media in foreign language learning and teaching: Subtitles for the deaf and hard-of-hearing as tools for language learning.* Springer.

Vanderplank, R., & Teng, M. F. (2024). Intentional vocabulary learning through captioned viewing: Comparing Vanderplank's 'cognitive-affective model' with Gesa and Miralpeix. In M. F. Teng, A. Kukulska-Hulme, & G. J. Wu (Eds.), *Theory and practice in vocabulary research in digital environments* (pp. 15–38). Routledge.

Vidal, K. (2003). Academic listening: A source of vocabulary acquisition?. *Applied Linguistics, 24,* 56–89. https://doi.org/10.1093/applin/24.1.56

Appendix A

Vocabulary Learning Strategies in Watching Audiovisual Materials

Instructions: Below are a number of statements about vocabulary learning. Please rate how much you personally agree or disagree with these statements. There is no right or wrong answer. We are interested in your ideas.

1	2	3	4	5	6	7
Strongly disagree	Moderately Disagree	Slightly disagree	Neutral	Slightly agree	Moderately Agree	Strongly agree

Emotion

1. I try to stay calm and focused when encountering difficult vocabulary words in audiovisual materials.
2. I take breaks when feeling overwhelmed or frustrated during vocabulary learning in audiovisual materials.
3. I remind myself that it's okay to not know every vocabulary word when watching audiovisual materials.
4. I use positive self-talk to boost my confidence in my vocabulary learning abilities when watching audiovisual materials.
5. I celebrate small successes in my vocabulary learning when watching audiovisual materials, such as learning a new word or using a word correctly in a sentence.
6. I visualize myself successfully mastering new vocabulary words when watching audiovisual materials.
7. I engage in mindfulness practices, such as deep breathing or meditation, to reduce stress and increase focus during vocabulary learning in audiovisual materials.
8. I seek support from peers, instructors, or other resources when feeling discouraged during vocabulary learning in audiovisual materials.
9. I identify any negative feelings related to vocabulary learning in audiovisual materials, then work to address them in a constructive way.
10. I manage my own emotions and stress levels during vocabulary learning in audiovisual materials to ensure a positive learning environment.
11. I feel relaxed to ask the instructor or other peers for clarification or examples of new vocabulary words from audiovisual materials.
12. I feel excited for watching English audiovisual materials.

Attention

13. I focus my attention on new vocabulary words when watching audiovisual materials.
14. I minimize distractions and interruptions to maintain my attention during vocabulary learning when watching audiovisual materials.
15. I use active listening and visual cues to maintain my attention during vocabulary learning when watching audiovisual materials.

16. I employ mindfulness techniques, such as deep breathing or meditation, to improve my focus and attention during vocabulary learning when watching audiovisual materials.
17. I adjust the playback speed or pause and rewind audiovisual materials to aid my comprehension and attention to new vocabulary words.
18. I take breaks when I feel my attention waning during vocabulary learning in audiovisual materials, then resume learning when I feel refreshed and focused.
19. I mentally repeat or rephrase new vocabulary words to help me maintain my attention and better remember the words.
20. I set specific goals for my vocabulary learning during watching audiovisual materials to keep me focused and motivated.
21. I use positive self-talk and affirmations to maintain a positive attitude and attention to vocabulary learning during watching audiovisual materials.
22. I engage in pre-reading or previewing activities, such as scanning headings or summaries, to help focus my attention on key vocabulary words before watching audiovisual materials.

Metacognition

23. I set specific vocabulary learning goals for myself when watching audiovisual materials, such as learning a certain number of new words per week.
24. I prioritize learning vocabulary words that are relevant to my field of study or interests when watching audiovisual materials.
25. I monitor my progress and adjust my learning strategies as needed when watching audiovisual materials to improve my vocabulary learning effectiveness.
26. I engage in self-testing or quizzing to reinforce my vocabulary learning after watching audiovisual materials.
27. I identify my strengths and weaknesses in vocabulary learning and use that knowledge to improve my strategies.
28. I monitor how complex vocabulary words can be broken into smaller parts or simpler concepts to help me understand and remember them when watching audiovisual materials.
29. I evaluate my notes or other study materials regularly to reinforce my vocabulary learning from audiovisual materials.
30. I reflect on my vocabulary learning progress after watching audiovisual materials to identify areas for improvement.

Information processing

31. I pay close attention to new vocabulary words when watching audiovisual materials.
32. I take notes on new vocabulary words when watching audiovisual materials.
33. I try to understand the context in which new vocabulary words are used when watching audiovisual materials.

34. I try to use new vocabulary words in sentences or discussions after watching audiovisual materials.
35. I try to connect new vocabulary words to previously learned material when watching audiovisual materials.
36. I try to use new vocabulary words in everyday conversations to reinforce the words learned from watching audiovisual materials.