



DOI:10.22144/ctujoisd.2026.029

EFL students' self-directed vocabulary learning through YouTube in a mobile learning ecology

Nguyen Thi Diem Thi^{1,2*}

¹Faculty of English, Hue University of Foreign Languages and International Studies, Viet Nam

²Faculty of Foreign Languages, Industrial University of Ho Chi Minh City, Viet Nam

*Corresponding author (ntdthi.dhnn23@hueuni.edu.vn)

Article info.

Received 29 Jul 2025
Revised 5 Oct 2025
Accepted 20 Feb 2026

Keywords

Activity Theory, mobile-assisted language learning, self-directed learning, vocabulary learning, YouTube

ABSTRACT

This study explores how Vietnamese EFL university students engage in self-directed vocabulary learning through YouTube. Guided by Activity Theory, it examines how YouTube functions as a mediational tool in EFL students' self-directed vocabulary learning and analyzes the key factors shaping students' engagement with YouTube. A total of 154 journals, 22 semi-structured interviews, and four focus group discussions were used as data for the study. Thematical analysis of these triangulated data uncovers seven significant learning patterns: (1) strategic and intentional, (2) incidental and affectively engaging, (3) personalized and interest-driven, (4) integrated with multiple digital tools, (5) repetitive for reinforcement, (6) influenced by educators and influencers, and (7) spontaneous and exploratory. Generally, it was found that YouTube served not merely as a content platform but as a dynamic mediating artifact in learners' evolving, self-regulated learning systems. This study emphasizes the efficacy of informal, self-directed vocabulary acquisition through platforms such as YouTube and contributes to discussions on digital literacy practices, learner agency, and integrated teaching frameworks.

1. INTRODUCTION

As learners can now engage with informal content outside the classroom, the shift from in-person language teaching to more informal resources via digital platforms such as YouTube and mobile-assisted language Learning (MALL) is gaining momentum (Fauziah & Diana, 2023). Anderson and Jiang (2018) noted that YouTube gained prominence in 2018 as the most favored online platform by 85 per cent of teens. Such settings offer authentic materials and flexibility, which improve the overall process of acquisition (Reinders & Benson, 2017). For English as a Foreign Language (EFL) learners, YouTube offers an extensive archive of user-generated, educational videos that holds great pedagogical value, particularly for enriching vocabulary knowledge.

The potential of YouTube in vocabulary teaching and learning has been documented in many studies. For example, YouTube can significantly enhance vocabulary acquisition and boost learner engagement, as EFL students often report that YouTube not only makes learning more enjoyable but also improves retention of new vocabulary (Kaboocha & Elyas, 2018). Class-based participatory action research projects also show that using YouTube videos increases motivation, active participation, and improved retention of target vocabulary (Heriyanto, 2015; Kaboocha & Elyas, 2018). Besides, Kaboocha and Elyas (2015) reported enhanced vocabulary acquisition among learners and improved engagement levels, necessitating the use of short clips in a Saudi EFL classroom. These benefits support the multimedia learning theory

whereby the rich audio and visual components on YouTube are more effective in strengthening comprehension and word mastery than text or audio, especially for younger learners (Dizon, 2022).

Despite the educational promise of YouTube and similar sites, there is a gap in research on the ways learners independently navigate these resources for language learning. Most MALL studies to date have focused on dedicated apps and structured programs, resulting in hundreds of publications on the effectiveness of language learning apps (Burston, 2015). By comparison, far fewer studies have examined learners' self-directed use of YouTube or other informal online media for improving language proficiency outside the classroom (Dizon, 2022). The limited studies include one by Benson (2015), which examined thousands of comments on English-language YouTube videos and concluded that approximately one in three comments demonstrated some form of language or cross-cultural learning through interaction. Recently, Bhusaery et al. (2024) and Musyaffa and Hapsari (2025) studied students' perceptions of YouTube as an informal learning medium for English vocabulary enhancement and concluded that the majority of students can benefit from watching YouTube videos to improve their vocabulary, and that YouTube also helps increase their learning motivation. This indicates that YouTube users are not merely passively entertained; they also engage in active meaning negotiation, language exchange, and peer learning through their expressive comments. However, the literature has yet to address how learners' participation on YouTube might fit within a larger, complex conceptual framework in Activity Theory (AT), highlighting the interrelations among individual approaches, mediational means, and contextual sociability. In addition, there is a knowledge gap around the independent use of YouTube by EFL tertiary learners in Vietnam for vocabulary learning, where English is a foreign language in higher education. This indicates a need for research that describes learners' strategies in this setting and situates their actions within a framework that addresses the cognitive, technological, and social dimensions of learning.

Therefore, this study aims to address this gap by examining EFL students' self-directed use of YouTube for vocabulary learning through the lens of AT. This framework focuses on the interconnected, mediated ecosystem surrounding learning, viewing it as a goal-oriented process with

many interacting parts, such as a subject (the learner), an object (the vocabulary or knowledge to be learned), digital mediators (like YouTube videos and mobile devices), the social community, and the relevant rules, divisions of labor, and structures pertaining to the activity (Hajimaghsoodi & Maftoon, 2020). As AT emphasizes the social aspects of learning and how the goals and motives of learners shape the use of instruments and content within a wider framework (Ellis, 2008), it enables the study to analyze self-directed YouTube usage by EFL students as an integrated activity system. Specifically, the research addresses the following questions: (1) How does YouTube function as a mediational tool in supporting EFL students' self-directed vocabulary learning?, and (2) What factors characterize students' engagement with YouTube for vocabulary learning? To provide a solid foundation for the study, the following sections review relevant literature.

1.1. Mobile-assisted vocabulary learning via media platforms

The concept of MALL refers to the use of mobile technologies in language learning and instruction, which helps make learning easier to access and allows students to study on their own (Kukulka-Hulme & Viberg, 2018). Within this framework, there are mobile language learning resources that provide learners with various ways to access, practice, and internalize vocabulary beyond the classroom. YouTube provides authentic, multimodal input through real-life content such as interviews, TED Talks, and instructional videos (Kelsen, 2009; Watkins & Wilkins, 2011). Its potential supports both intentional and incidental vocabulary learning thanks to its audio-visual components as well as the automated subtitles (Musyaffa & Hapsari, 2025) (automatically generated or user-uploaded) and contextual recommendations. However, it is a question for future studies to scrutinize the accuracy of the subtitles in cases where language learners use them. Despite extensive scholarship on MALL and video-based learning, the intersection of YouTube and self-directed vocabulary learning is comparatively underexplored.

1.2. Self-directed learning (SDL) and digital autonomy

SDL is characterized by learners taking initiative in setting goals, selecting resources, applying strategies, and monitoring progress (Garrison, 1997; Knowles, 1975), which, in the digital era, is

increasingly intertwined with learners' ability to exploit online environments. For example, Lai and Zheng (2018) noted that mobile learners are more likely to adopt a more actively explorative approach by seeking and adjusting input based on its relevance and their engagement. Digital autonomy, however, varies significantly among individuals as they differ in their ability to filter, focus, sustain attention, and execute media-rich vocabulary tasks (Reinders & Benson, 2017).

With the use of YouTube as a learning platform, the variance is expanded with students' engagement levels in structured vocabulary practice (e.g., deliberate note-taking, rewatching) or incidental exposure (e.g., passive entertainment viewing). Thus, effective SDL through media like YouTube not only requires motivation and strategy, but also learners' metacognitive thinking regarding how to convert media input into language outcomes.

1.3. Incidental and intentional vocabulary learning via videos

Vocabulary, especially receptive vocabulary acquired through listening and reading, is critical for language mastery and flourishing academically (Nation, 2013). As noted by Hulstijn (2003), vocabulary acquisition can occur both through intentional means, like dedicated memorization and testing, as well as through unintentional exposure where learners engage with the content. The literature also indicates that learners' retention and mastery of word knowledge are enhanced by repeated contextual exposure to new words (Schmitt, 2008; Webb, 2007). For example, video materials, especially those that are captioned, aid in vocabulary development by reinforcing the association between words and their meanings, as well as through contextual use (Montero Perez et al., 2014). The unique feature of YouTube as both a source of entertainment and an educational tool allows learners to switch between intentional and incidental modes of learning. Nevertheless, cognitive overload, distraction, and engagement with content that does not match the learners' skill level may result from a non-linear structure (Kalyuga et al., 2003; DeStefano & LeFevre, 2007; Park et al., 2011; Ward et al., 2017).

As noted in Zhang and Zou (2020), learners can transform informal content into episodes of vocabulary learning by integrating reflection, annotation, and even digital flashcards and dictionaries into YouTube videos. However, more research is needed to understand how learners

maneuver through informal settings, how these patterns of use aid in vocabulary retention, and which patterns are most helpful for long-term retention.

1.4. AT and tool-mediated learning

To explore the impact of YouTube on vocabulary acquisition, this study adopts AT (Engeström, 1999, 2015) as its conceptual framework. The theory approaches human action as purposeful, within social contexts, and it incorporates a dynamic system of subject, tools, object, rules, community, and division of labor. Engeström's expanded activity system also allows researchers to examine contradictions (e.g., distraction vs learning focus), tool appropriation, and expansive learning, in which learners restructure their activity to overcome limitations. Within this framework, YouTube is viewed as a mediating artifact that both supports and limits the learners' efforts towards vocabulary development. and learners, subjects of the activity, adapt and transform the tool through iterative interaction. By that means, their behaviors are shaped not only by cognitive goals but also by social and technological contexts. Thus, in the context of mobile learning, AT provides a holistic lens for understanding how learners integrate various forms of media, planning, self-regulation, and strategies into the vocabulary learning process.

2. METHOD

2.1. Research design

This study utilized a qualitative case study design to investigate the ways in which EFL university students utilized YouTube as a mobile resource for SDL of vocabulary. The case study approach was effective in documenting the situated experiences, actions, and reflections of learners as they interacted with digital tools outside the classroom. AT was used as the interpretive framework to structure data collection, coding, and analysis by focusing on YouTube's mediational role within the learners' activity systems.

2.2. Participants

The participants were undergraduate students studying English at a Vietnamese university, where English has been taught as a foreign language over a four-year curriculum. A total of 24 students who confirmed having conducted SDL of vocabulary via mobile learning resources were purposively selected, with three students from each academic level (freshmen, sophomores, juniors, and seniors)

to capture a variety of viewpoints from learners at different educational levels. Of these 24 students, 22 documented their experiences through weekly learning journals and participated in individual interviews, and 19 participated in the focus group discussions. The participants were all native Vietnamese from different geographical locations, and were 18 to 22 years of age. Of the 24 students, 23 were female, which is in line with the demographics of such English language programs. Their English proficiency varied considerably but generally ranged from intermediate (B1) to upper-intermediate (B2). Their proficiency was influenced by prior exposure to English, motivation, learning strategies, levels, and the accessibility of YouTube as a mobile learning resource. All participants voluntarily self-reported using YouTube for vocabulary learning as part of their out-of-class English study routines. To maintain participant anonymity, pseudonyms were used alongside alphanumeric codes (e.g., S2.1.1).

2.3. Data collection

Three data sources were used for triangulation purposes. For the first source, the participants maintained weekly learning journals for seven weeks, documenting their vocabulary learning activities. Each week, they noted the learning motivations, the strategies they executed, and their self-evaluation regarding the efficacy of their learning, which facilitated the appraisal of progress and changes in behavior over time. The journals also documented vocabulary knowledge learned in terms of meanings, pronunciation, collocations, and contextual use, as well as the methods students used to retain the words. Students finally reported the learning outcomes in terms of the ability to recall, or put vocabulary into use. The second part of the research involved semi-structured individual interviews ($n = 22$), each lasting 20-30 minutes and conducted in Vietnamese. These interviews examined students' motivations, strategies, and perceptions regarding YouTube. After that, nineteen out of the 22 students agreed to join focus group discussions which looked at collective behaviors, issues, and peer influences about SDL of vocabulary through YouTube. They were divided into four groups based on level of study, i.e., freshmen/sophomores and juniors/seniors, securing that views were organized to compare distinct stages of university. This arrangement enabled the study to identify possible differences not only between inexperienced and experienced students, but also between students who are more and less involved in

SDL. All qualitative data were transcribed, translated into English, and checked for meaning and accuracy through several back-checking procedures such as transcript verification, expert review, back-translation and participant validation.

2.4. Data analysis

The data was manually analyzed thematically based on AT using a coding framework within a three-step iterative process. To strengthen the validity of the findings, two independent coders reviewed 20% the dataset, for which Cohen's kappa was computed and reached 0.82 signifying good agreement. In the event of discrepancies, discussions were held to reach agreement, and the coding scheme was implemented across the full dataset once consensus was reached. Comprehensiveness was ensured when data saturation was reached, specifically when the last two focus groups and interview transcripts were analyzed, and no new codes or patterns emerged.

Operationalized constructs of the Activity Theory were outlined in the coding framework: Subject – the students engaging with YouTube; Mediating tool: YouTube and other mobile learning resources (dictionaries, apps, AI tools); Object: the acquisition and retention of vocabulary knowledge; Rules: norms and expectations which shape learning (e.g. institutional expectations, self-imposed routines); Community: peers, influencers, and teachers who interact with learners or shape practices of the learners; Division of labor: the roles and responsibilities of learners, peers, teachers, and influencers; and Contradictions: tensions and mismatches learners face (e.g. enjoyment and distraction, control vs autonomy, informal learning vs formal expectations). These tensions were captured as contradictions driving change in the learning activity system.

Initially, open coding was performed to capture all segments associated with YouTube utilization, vocabulary strategy execution, tool-related actions, and learner self-assessments. Then, axial coding synthesized these segments into thematic categories aligned with the major components of AT: subject, tool, object, and contradictions. Finally, thematic synthesis was used to develop broader analytical themes, including intentional learning, incidental learning, tool integration, repetition, and learner autonomy, as illustrated in Figure 1 below. This approach facilitated the identification of both cognitive dimensions (e.g., strategy use) and contextual elements (e.g., motivation, digital ecosystems) within learners' activity systems.

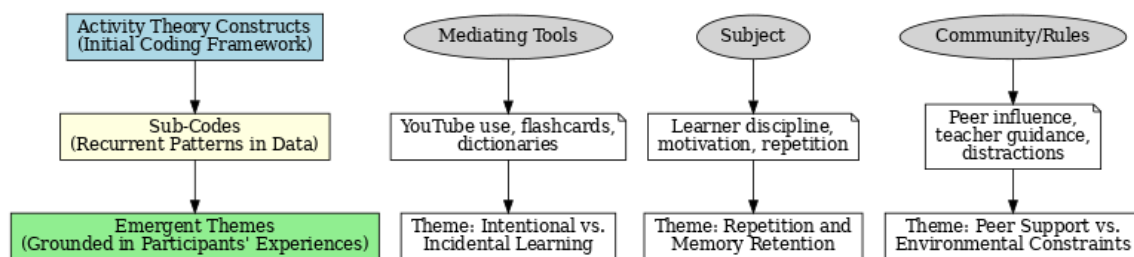


Figure 1. Diagram of theme generation process

Figure 1 indicates that AT was the overarching framework for data analysis. However, some themes, such as intentional vs incidental learning or repetition, emerged from the data inductively as the participants described their experience of vocabulary learning. For instance, while working on the tools (YouTube) and the object (vocabulary learning), the students' reflections noted differences between deliberate practice and incidental acquisition of vocabulary during authentic media use. Under rules and division of labor, they also described vocabulary being retained through repeated exposure and the repetition as a strategy for retaining vocabularies. These themes can be seen as sub-patterns within the larger categories in AT and add explanatory depth to the framework.

To establish credibility, trustworthiness, and member verification, the research findings were triangulated across the three data sources. Interpretation verification checks were performed with the participants themselves. Researcher reflexivity was also maintained throughout the study to mitigate potential bias in data coding and interpretation.

2.5. Ethical considerations

The study was conducted in alignment with the ethical principles established by the university. There was no ethics review board in place at the time of research, but the ethics of the research were nonetheless fully respected. Each participant was informed of the purpose of the research, the procedures, and their rights, including the right to withdraw at any time without consequences. Before taking part in the study, participants provided their written informed consent. To ensure confidentiality, all participants were assigned pseudonyms and alphanumeric codes in lieu of their real names.

Other personal information and contacts were kept secure and accessible only to the researcher. All data was kept secure and used for academic purposes only. Potential risks to the participants were kept minimized.

3. RESULTS AND DISCUSSION

3.1. Results

The themes of the study were constructed through multiple readings of the qualitative data, initially guided by the principles of AT but subsequently refined through an inductive approach based on the participants' narratives. While AT constructs such as the subject, tools, objects, and rules provided initial scaffolding for the coding process, the students' reflections revealed more complex themes that went well beyond these constructs. These themes illustrate the points the participants had reached in their self-directed vocabulary learning and document the YouTube and mobile resources. In the following section, each theme is presented with supporting evidence from interviews, journals, and focus group discussions.

3.1.1. Strategic and intentional vocabulary learning via YouTube

A significant number of students reported intentionally using YouTube to enhance their vocabulary. Their viewing strategies incorporated watching specialized sessions such as TED Talks, taking pauses to note new words, replaying sections of videos to reinforce vocabulary, and enabling subtitles for better comprehension. These behaviors exemplify tool-mediated learning since learners actively repurpose YouTube from a passive entertainment resource into an interactive and personalized education environment.

Table 1. Analysis of Theme 1: Students' strategic and intentional vocabulary learning via YouTube

Code / Sub-theme	Illustrative quotation
Watching strategy-sharing content	<i>"I watch videos on YouTube on how to learn vocabulary, they share their experiences, and I learn from them."</i> - S4.1.3 (Journal)
Using subtitles and AI tools for accuracy	<i>"I use YouTube subtitles to cross-check what I understand with the actual content and use AI to make flashcards."</i> - S2.1.1 (Interview)
Following a structured learning sequence	<i>"I learn vocabulary by looking up definitions, listening and practicing pronunciation, writing them down, and noting examples of the words in context. Then I search for YouTube videos that include those words."</i> - S2.2.2 (Journal)
Using TED Talks and advanced input intentionally	<i>"I also consider watching videos from the TED YouTube channel as a fun way to learn new words at an advanced level."</i> - S2.2.3 (Interview)
Collaborative vocabulary learning with peers	<i>"I searched YouTube myself and invited friends to join me to study, which was quite effective for group learning."</i> - S2.1.1 (Interview)
Note-taking and dictionary follow-up routines	<i>"I watch videos on YouTube, take notes of words I don't know in my notebook, and look them up in online dictionaries later..."</i> - S4.1.3 (Journal)
Cross-platform strategy: YouTube and flashcards	<i>"Use Quizlet, study with flashcards, and watch YouTube videos on the topic being learned."</i> - S2.2.1 (Interview)

For example, one student shared that they often searched for vocabulary learning strategies on YouTube and learned by watching others explain their methods: *"I watch videos on YouTube on how to learn vocabulary, they share their experiences and I learn from them"* (S4.1.3, Journal Entry). Another reported monitoring comprehension using YouTube's subtitle feature and subsequently creating flashcards with AI tools such as ChatGPT (S2.1.1, Interview). These response patterns demonstrate how learners adapt the features of YouTube to serve their individual objectives, in this instance, vocabulary acquisition, which is consistent with the principles of AT and its emphasis on agentive engagement with the tool.

Conversely, other participants regarded YouTube as a supplementary resource for revisiting previously learned vocabulary. One participant explained their approach started with looking up definitions, listening to the words, practicing their pronunciation, crafting example sentences, and then searching YouTube videos that were using the words in context (S2.2.2, Journal). Besides, the participants also showed purposeful selection of content related to the cognitive demand they perceived. One of them suggested that they found TED YouTube channel videos enjoyable and useful for developing advanced vocabulary knowledge (S2.2.3, Interview). These examples demonstrate consistent self-regulated learning strategies and illustrate that learners actively adjust their

preferences to include challenging materials aligned with their objectives.

Along with personalized routines, the learners also created collaborative routines of their own. They kept searching for appropriate content on YouTube and afterward held study groups, which, in their words, were "quite effective" (S2.1.1, Interview). This captures a more comprehensive perspective on strategy use that considers social learning as a form of SDL, thereby expanding the boundaries of individual practices. Additionally, the students integrated watching videos with vocabulary notebooks, digital dictionaries, and even flashcard applications. For example, one participant stated, *"I watch videos on YouTube, take notes of words I don't know in my notebook, and look them up in online dictionaries later"* (S4.1.3, Journal Entry). Another also described a triadic routine of using Quizlet, reviewing flashcards, and watching YouTube videos directly related to their topics of study (S2.2.1, Interview). The integration of YouTube with other tools reveals how learners use YouTube within wider self-organized vocabulary consolidation and review frameworks.

Overall, these results underscore the multiplicity of strategic, purposeful, and goal-oriented actions by learners to leverage YouTube as a flexible, multimodal educational tool. Their efforts indicate not only cognitive investment but also metacognitive frameworks for planning, tool integration, and consolidation, which are

characteristic of independent learning. Aligned with AT, learners are not passive consumers in this study; they are active participants who dynamically adapt their use of technological resources to transform and fulfill shifting content and context-driven learning requirements.

3.1.2. Incidental and enjoyable learning

In addition to the goal-directed approaches

employed by many learners, some participants noted that they sought vocabulary knowledge while casually browsing YouTube. These spontaneous learning opportunities often involved an emotionally engaging context that, without a definitive study goal, appeared to support vocabulary retention.

Table 2. Analysis of Theme 2 – Incidental and enjoyable vocabulary learning via YouTube

Code / Sub-theme	Illustrative quotation
Vocabulary acquisition through emotional content	<i>"I learn more unintentionally. When I enjoy the material, I can remember words better. Intentional learning feels dry, so it's harder to remember."</i> - S3.1.2 (Focus Group)
Passive viewing evolving into active learning	<i>"Learning through YouTube, taking notes, practicing pronunciation, then writing a paragraph."</i> - S2.1.1 (Interview)
Impact of device and viewing context	<i>"Learning via YouTube on a computer is more effective than on a phone. Situational visuals increase motivation."</i> - S3.2.2 (Interview)

From the data, it was reflected that vocabulary learning felt more effective when it occurred during enjoyable video consumption,

"I learn more unintentionally. When I enjoy the material, I can remember words better. Intentional learning feels dry, so it's harder to remember. For instance, I learn more from watching YouTube videos on various topics. I listen to news more than I read it, I enjoy listening more" (S3.1.2, Focus Group).

This example illustrates the interplay between incidental and intentional learning and shows how an emotional connection can enhance retention and the desire to participate actively. Another participant noted the value of emotional involvement in content, as it enhanced focus and vocabulary retrieval. As one student described their learning as evolving from passive to more active levels: *"Learning through YouTube, taking notes, practicing pronunciation, then writing a paragraph"* (S2.1.1, Interview), it can be realized that learners may eventually shift into performing some sort of activity and become more focused on that one, which demonstrates a flexible framework of informal and mobile learning via YouTube. Besides, how students accessed learning content also shaped the learning experience. Learner S3.2.2's preference for learning via YouTube on a

computer illustrates the idea that the physical context and the tools used to interact with the content are supplementary to their cognitive engagement, even in cases of incidental learning.

The examples illustrate what AT refers to as emergent object construction. In this case, learners do not set out to develop vocabulary, but over time, as they interact with digital media, vocabulary acquisition becomes a focus. The study also recognized that incidental learning is not entirely random. Instead, it arises from learners' interests, emotions, and media use in richly mediated environments, and YouTube is an excellent example of such environments, as its diverse genres and algorithmically curated recommendations provide critical language input in enjoyable, relatable settings, fostering spontaneous yet powerful learning.

3.1.3. Personalized and interest-based content access

YouTube provided students with an extensive range of genres and subjects as well as opportunities to access materials aligned with their personal interests and academic goals. Students frequently reported that unrestricted exploration led to heightened motivation, sustained attention, and even greater vocabulary retention.

Table 3. Analysis of Theme 3 – Personalized and interest-based content

Code / Sub-Theme	Illustrative Quotation
Extended engagement with preferred topics	<i>“My level is between B2 and C1... I spend about 2 hours in the evening watching videos on YouTube on topics I enjoy.” - S3.1.3 (Journal)</i>
Motivation through interest-driven content	<i>“I spend about 2 hours in the evening watching videos on YouTube on topics I enjoy.” - S3.1.3 (Journal)</i>
Using TED Talks for vocabulary enhancement	<i>“Using videos from the TED YouTube channel can enhance my vocabulary knowledge and identify topics in the videos I watch.” - S4.1.1 (Interview)</i>
Personal selection of cognitively rich content	<i>“Using videos from the TED YouTube channel can enhance my vocabulary knowledge...” - S4.2.3 (Interview)</i>

Multiple respondents highlighted the role of their interests in selecting content. For instance, as one participant depicted, *“My level is between B2 and C1... I spend about 2 hours in the evening watching videos on YouTube on topics I enjoy”* (S3.1.3, Journal), it can be realized that when the content offered by the materials aligns with their curiosity or entertainment preferences, learners are more willing to invest substantial time to content engagement, thus facilitating meaningful vocabulary input. Besides, participants also described TED and similar sources to be intellectually stimulating and rich in language, providing access to advanced vocabulary in authentic contexts. In interviews with learners S4.1.1 and S4.2.3, they confirmed that they use the TED YouTube channel to find video topics to watch to enhance their vocabulary. This fact demonstrates that the category of selected content also had a notable influence on perceived learning benefits, highlighting the importance of learner interest for vocabulary development.

As per AT, learners' choice of tools and content is determined by their motives and objectives for learning. In this scenario, YouTube functioned as a mediational artifact that enabled learners to connect

their personal interests with the object of learning vocabulary, thereby supporting sustained investment towards exposure to English beyond the rigid temporality of the classroom. Students' comments, therefore, illustrate that having personalized access to content not only increases their enjoyment but also enhances their ownership of the learning process. The combination of digital recommendation systems and learner agency provides a fluid framework in which vocabulary acquisition develops as a byproduct of ongoing interaction with rich, relevant materials.

3.1.4. Tool integration and learning ecosystem

Most students did not restrict themselves to YouTube. Instead, they integrated it into a wider network of digital resources, forming complex self-regulated learning ecosystems. These systems included dictionaries (Oxford, TFlat), flashcard apps (Quizlet, AnkiDroid), AI tools (ChatGPT), and other media such as podcasts, music, and educational programs. Learners used these tools not only for vocabulary exposure but also for clarification, review, production, and self-assessment.

Table 4. Analysis of Theme 4 – Tool integration and learning ecosystem

Code / Sub-Theme	Illustrative Quotation
Combining YouTube with AI and note-taking	<i>“I use ChatGPT and watch videos on YouTube. Then I take notes, watch videos, and create vocabulary lists.” - S2.1.1 (Interview)</i>
Parallel use of dictionaries and mobile apps	<i>“I also use Oxford... I watch YouTube videos and TikTok with my phone.” - S4.2.1 (Focus Group)</i>
Strategy transfer from YouTube to routines	<i>“If I don't know a word, I guess its meaning from the context first... I learned this method from YouTube, copied it, and then adjusted it.” - S2.2.2 (Interview)</i>
Extensive digital toolset with YouTube hub	<i>“YouTube podcasts, Google Translate, ChatGPT, TFlat, and Quizlet.” - S4.2.2 (Interview)</i>

For example, one student described this multi-tool routine as follows: *“I use ChatGPT and watch videos on YouTube. Then I take notes, watch videos, and create vocabulary lists”* (S2.1.1, Interview). This quotation captures the learners’ use of synchronous and asynchronous cross-platform tools, which enable them to construct a tailored and iterative process for vocabulary acquisition. This combination not only demonstrates diverse input modalities but also intentionally blends cognitive processes and motivational dimensions. Apart from using tools, the learners demonstrated substantial metacognitive awareness in modifying strategies across different platforms. One student provided an example: *“If I don’t know a word, I guess its meaning from the context first, then look it up on Google... I learned this method from YouTube, copied it, and then adjusted it”* (S2.2.2, Interview). This illustrates strategy transfer, in which learners appropriate techniques from informal learning experiences and seek to modify them through personal initiative. Students’ lists of various tools they use in their SDL of vocabulary, besides YouTube, suggest that the ecosystem is dynamic rather than static and changes as a result of reflection. This willingness to iterate indicates the possibility of learner agency for building and networking tools that are personally applicable.

Social learning contexts were also advanced by some individuals through recommending and sharing digital tools with others, illustrating a collaborative aspect of tool integration whereby learners function as facilitators of other learners’ resource-driven learning. Finally, the scope of tools integrated with YouTube was particularly wide. Students’ lists of tools, such as “YouTube podcasts, Google Translate, ChatGPT, TFlat, and Quizlet” (S4.2.2, Interview), reinforce YouTube’s central position in the multimodal learning network. Through this ecosystem-based approach, learners created layered learning environments that

supported vocabulary input, meaning construction, and knowledge consolidation. Informed by AT, this evidence highlights that learners’ tool selection is influenced by their goals, the learning object (in this case, vocabulary), and the learning environment. The use of YouTube was not in isolation as a resource. Rather, it was a mediating artifact within the activity system of learning.

3.1.5. Repetition and reinforcement

An important theme that emerged from the student reflections concerned the importance of repetition in reinforcing already acquired vocabulary. This perspective was particularly salient given that students claimed to typify non-linear, consumption-oriented review of materials. They deployed purposeful reviews that transformed incidental input into knowledge that could be retrieved, which reflects learners’ efforts to recycle materials. For instance, a number of students reported intentionally rewatching videos as part of their study regimen. One participant described, *“I replay those videos and practice using flashcards”* (S2.1.1, Focus Group), indicating a synthesis of passive review and active recall through flashcard use. Another stated, *“Maybe I need to see them several times to memorize them”* (S4.2.1, Focus Group). This statement shows understanding of the cognitive effort that vocabulary memorization entails, along with the need for spaced repetition.

Learners also appeared to build automated systems to aid in the reviewing processes, utilizing digital methods. One participant put it, *“In my free time, I watch videos on YouTube or TikTok... if I find new words, I often use TFlat to look them up so that I can be reminded of them by the app the next day”* (S4.2.1, Interview). This quote exemplifies how learners go beyond passive content consumption. They create customized systems consisting of exposure, lookup, and timed review, which incorporates repetition into the review process.

Table 5. Analysis of Theme 5 – Repetition and reinforcement

Code / Sub-theme	Illustrative Quotation
Rewatching for retention	<i>“I replay those videos and practice using flashcards.”</i> - S2.1.1 (Focus Group)
Multiple exposures for memorization	<i>“Maybe I need to see them (the vocabulary) several times to memorize them.”</i> - S4.2.1 (Focus Group)
Tool-supported review loops	<i>“In my free time, I watch videos on YouTube or TikTok... I often use TFlat to look them up so that I can be reminded of them by the app the next day.”</i> - S4.2.1 (Interview)

The provided examples show that despite the mobile learning environment's seemingly unstructured character, learners have control over the times and ways they interact with language input. Analyzed through the framework of AT, these behaviors illustrate the learner's agentic adaptation of mediating tools and artifacts in pursuit of an enduring learning objective. Students actively cope with vocabulary acquisition by designing personalized review cycles that include flashcards, automated reminders, and media repetition, demonstrating the long-term potential for digital tools in vocabulary enhancement.

3.1.6. Peer and influencer support

Learners emphasized the significant role English vocabulary content creators played in their development of English vocabulary skills. These individuals acted as trustworthy suppliers of

language input, learning strategies, and motivational feedback. To some students, YouTubers and other online educators seemed to connect with them better than their peers or even their teachers; as one participant observed, *"Some YouTubers or influencers are more helpful than classmates"* (S4.1.1, Focus Group). This explains why online creators have become preferred models for learning because of their consistent style, custom-tailored content, and ease of access. Another participant reflected, *"I found a teacher on YouTube that really suits me, called The Forum"* (S2.1.2, Journal), indicating the emergence of a parasocial connection with instructors online, which demonstrates a shift from peer-to-peer assistance toward learner-influencer associations through social network platforms rather than face-to-face interaction.

Table 6. Analysis of Theme 6 – Peer and influencer support

Code / Sub-theme	Illustrative quotation
Influencers more impactful than peers	<i>"Some YouTubers or influencers are more helpful than classmates."</i> - S4.1.1 (Focus Group)
Personal alignment with digital educators	<i>"I found a teacher on YouTube that really suits me, called The Forum."</i> - S2.1.2 (Journal)
Multimodal access for influencer content	<i>"I use a phone (to watch lessons on YouTube) and a laptop."</i> - S3.2.2 (Journal)
Shift away from teacher dependence	<i>"I learned these methods through online video tutorials. Teachers only told me to copy vocabulary..."</i> - S3.2.2 (Journal)

From the perspective of AT, these influencers and mobile learning tools form part of the learners' broader community and division of labor, elements that shape learning outcomes through social, cultural, and technological mediation. Through social media influencers, learners gain access to vocabulary and strategies while developing SDL through independent learning behaviors, which boost their motivation and competence outside formal school settings. These results, in turn, provide evidence of the extent to which cultures of digital learning are transforming traditional support frameworks. Online educators and peer-shared content are more than added value; they are indispensable to learners on the way to accessing, internalizing, and replicating vocabulary learning behaviors in an autonomous system.

3.1.7. Spontaneous and exploratory engagement

In addition to planned and organized learning, many students mentioned engaging with YouTube and related platforms in a more spontaneous way. This type of exploratory learning is more likely to occur

through mobile phone use, algorithmic suggestions, or spontaneous learning rather than through set goals. Despite the lack of formal structure, such engagement can also lead to incidental vocabulary acquisition and reflective adaptation over time.

One student clearly described this unstructured approach: *"I don't have specific plans, just do it unintentionally, randomly"* (S4.1.3, Journal). Another explained their learning to be *"Mostly spontaneous. Sometimes we align with class topics, but I usually just scroll TikTok or YouTube and learn randomly"* (S4.1.2, Focus Group). These examples illustrate how learners allow their media environments to guide learning encounters and indicate a hybrid model in which formal and informal goals occasionally intersect, with spontaneity predominating. In addition, their ability to shape emergent learning pathways was also emphasized, according to a student's comment, *"Videos on YouTube and TikTok are accessed spontaneously. I find these resources on my own..."* (S4.2.3, Interview). Learners were taking advantage

of digital freedom to seek out materials that suited their interests or language skills at that moment, illustrating AT's focus on emergent object formation, as the purposes of a particular learning activity change in relation to the available resources and the surrounding context. Besides, another noted that "As for videos on TikTok or YouTube... I can

save them and next time I can watch again to review words" (S4.2.3, Interview), illustrating that learners tend to purposefully recapture content in subsequent engagements once they realize its potential for learning, even if the initial encounter was unintentional.

Table 7. Analysis of Theme 7 – Spontaneous and exploratory engagement

Code / Sub-theme	Illustrative quotation
Random mobile browsing behavior	<i>"I don't have specific plans, just do it unintentionally, randomly."</i> - S4.1.3 (Journal)
Algorithm-driven and self-guided exploration	<i>"Videos on YouTube and TikTok are accessed spontaneously. I find these resources on my own..."</i> - S4.2.3 (Interview)
Hybrid spontaneity with occasional alignment	<i>"Mostly spontaneous. Sometimes we align with class topics, but I usually just scroll TikTok or YouTube and learn randomly."</i> - S4.1.2 (Focus Group)
Emergent review strategies post-discovery	<i>"As for videos on TikTok or YouTube... I can save them and next time I can watch again to review words."</i> - S4.2.3 (Interview)

In sum, these descriptions are examples of adaptive agency, a specific form of learner behavior that is responsive and characterized by flexibility and curiosity toward media ecosystems. Although unintentional spontaneity lacks a formal structure, it facilitates vocabulary acquisition through meaningful, emotionally resonant, and repetitive exposure to English input. From the lens of AT, these patterns illustrate how learners collaboratively interpret and actively transform their learning processes through sustained interaction with digital tools and spaces.

3.2. Discussion

This section discusses the findings according to the two objectives of the study: (1) understanding the mediational role of YouTube in EFL students' self-directed vocabulary learning and (2) examining the factors that define their engagement with YouTube in learning vocabulary. In answer to the first objective (RQ1), the mediational role YouTube played in learners appropriation of the tool ecosystems and integration with other resources to aid in vocabulary learning is illustrated in the themes of personalization and autonomy in content choice. As for the second objective (RQ2), the themes of strategic and intentional learning, incidental and enjoyable learning, repetition and memory, influencer learning, and spontaneous engagement reflect multifaceted factors that defined students' engagement with YouTube. These factors illustrate the cognitive, technological, social, and affective dimensions in students' vocabulary learning and development via YouTube.

3.2.1. Personalization and autonomy in content selection

Section 3.1.3 has demonstrated how personalization was important to students' continued engagement with English vocabulary learning via YouTube. It is evident that learners were not passive in response to algorithmic suggestions; instead, they actively selected content tailored to their preferences and learning objectives. This finding is consistent with earlier research on the motivational effects of passion-driven learning in informal digital environments (Reinders & White, 2016). Participants' active engagement with self-selected content, such as TED Talks and other videos, demonstrated the extent to which autonomy in content selection boosts cognitive engagement, active learning, and vocabulary retention. In this case, AT helps explain these personalized learning pathways and micro-scenarios focused on SDL. The learner (subject) picks a certain content (mediating tool) that motivates them. The object is transformed from passive watching to actively seeking ways to enhance vocabulary. This transformation results from the balance of internal motivation and external factors – YouTube's algorithmic suggestions.

These examples demonstrate how learners make informal contexts more structured and relevant to their needs. By choosing advanced-level inputs such as TED Talks, learners transform from passive media consumers to active participants in language exposure. This aligns with Lai and Zheng's (2018) study on digital settings where personalization augments learner agency and goes beyond traditional learning ecologies, echoing Benson

(2011), who referred to learning outside the classroom where students collaboratively gather materials beyond textbooks. The combination of learners' self-directed efforts and YouTube's features exemplifies progress toward flexible, autonomous educational frameworks. Rather than relying on rigid institutional paradigms or resources predetermined by instructors, learners select their materials by themselves, which reinforces their autonomy and control and actively motivates long-term engagement with vocabulary learning. The balance between learner agency and the platform's affordances illustrates that purposeful informal vocabulary learning occurs when content is curated and contemplated.

3.2.2. *Tool ecosystems and expansive learning in vocabulary development*

The findings in Section 3.1.4 illustrate the self-regulated nature of learners' SDL of vocabulary and the significant role that YouTube, as part of a larger, more complex digital learning ecosystem, plays. This correlates with the overarching findings of self-regulated learning with technology and is particularly devoted to how learners use different tools to learn new vocabulary at various levels - input, clarification, practice, and review (Lai & Gu, 2011; Reinders & White, 2016). Based on AT, learners' tool integration exemplifies expansive learning; an activity system is persistently transformed as new mediational artifacts are integrated (Engeström, 2015). Participants described tailored workflows that integrated YouTube, dictionaries, flashcard apps, AI tools like ChatGPT, and even social media, where each tool offered distinct advantages for meaning-making, self-evaluation, or collaborative benchmarking. This ecosystem perspective is consistent with Lai and Gu's (2011) self-regulated multi-tool learning model, which depicts a self-directed blend of input, practice, review, and reflection with various tools.

Moreover, the socio-collaborative elements further enriched the learning ecosystem. Some participants not only shared specific vocabulary-application and content platforms but also took on active mentoring roles, demonstrating that learners did not merely act as passive resource consumers but rather circulated resources within their networks. Reinders (2014) described personal learning environments as systems where learners self-manage their learning and assist others in improving their digital competencies. The integration of AI technologies like ChatGPT into these frameworks indicates a

move towards an AI-augmented pedagogical framework as pointed out by Zhang and Zou (2020). This indicates a shift towards a more sophisticated version of community in AT, where the idea of peers has expanded beyond classmates to include digital contemporaries, AI, and networks outside of school.

As a whole, these findings reinforce the idea that vocabulary acquisition in mobile settings depends not on the tools used but on the learning environment. YouTube acted as a gateway to a wider array of resources that learners coordinated with increasing sophistication and autonomy. By situating vocabulary learning within a constellation of tools and social practices, students exercised control over their own learning environments, transforming YouTube from a single mediational artifact into a personalized, strategic, and collaborative learning system.

Discussion of research question 1

From the discussions of the two themes – personalization and autonomy in content selection and tool ecosystems and expansive learning – the findings indicate that YouTube was more than a provider of algorithmic input; it could be viewed as a place where learners self-curated goal-aligned content, thereby demonstrating autonomy in carving their own pathways. Furthermore, YouTube was part of a broader ecosystem that included dictionaries, flashcards, AI tools, and peers, all of which supported expansive learning. The learners were able to transform YouTube, typically viewed as passive entertainment, into an active mediational artifact that facilitated meaning-making, practice, and reflection. In this way, YouTube served as both a primary entry point for authentic vocabulary input and a gateway to myriad other tools, thereby mediating self-directed vocabulary learning in a flexible and dynamic manner.

3.2.3. *Strategic and intentional vocabulary learning via YouTube*

Under this theme, the findings emphasize the learners' efforts in controlling the use of YouTube as a vocabulary learning tool. Participants did not use the platform as a passive resource; instead, they took considerable measures, such as engaging with appropriate materials, arranging the sequence of learning activities, and even employing several requisite instruments to transform casual video viewing into learning activities. Such actions align with the concept of tool mediation within the framework of AT, where learners strategically

revise their use of digital tools for vocabulary acquisition (Engeström, 2015).

The data suggest that learners did not simply consume content passively but actively engaged in content curation. For example, some students selected videos, such as TED Talks, to address their linguistic and cognitive challenges. Others were engaged in a learning process that began with pronouncing words from a dictionary, followed by searches on YouTube for the contexts of the newly learned words. This sequencing demonstrates metacognitive planning and self-regulation, key components of SDL (Garrison, 1997; Knowles, 1975). Besides, a noteworthy dimension is the multi-platform strategy employed by learners. Rather than relying on YouTube in isolation, students integrated it with tools such as Quizlet, TFlat, and ChatGPT, forming personal learning ecosystems (Reinders & Benson, 2017). This is consistent with Sydorenko (2010) and Dizon's (2022) emphasis on the need for multi-layered methods, and with Lai and Zheng's (2018) remark that in digital environments, learners self-customize skill sets that reflect their learning styles, objectives, and levels of digital literacy. In this study, students customized their learning ecosystems by integrating and interacting with content input through YouTube, meaning-making with dictionaries and AI, and reinforcement via flashcards, thus achieving repetition and meaningful interaction with lexical items.

Some self-directed learners expanded their practices by inviting peers to collaborate through YouTube sessions or by sharing resources. This illustrates the rise of social strategy and the community aspect in AT, which views learning not as an individual endeavor but as mediated through social tools and constructs. Additionally, the explicit use of YouTube to learn how to learn vocabulary, where they watched meta-strategic content produced by influencers or educators, emphasizes the reflexive aspect of learner agency. This illustrates a shift from content acquisition to learning strategy acquisition, a form of what Oxford (2011) characterizes as higher-order strategy use.

To conclude, learners' abilities to navigate YouTube go beyond seeking English-language exposure to actively constructing their learning journey. These patterns of behavior demonstrate expansive learning, in which learners seek to address an identified gap, such as the absence of scaffolding, by adapting their context, instruments,

and practices. Such behaviors also support the view of learners as active agents capable of designing personalized, goal-oriented vocabulary learning pathways within informal digital spaces.

3.2.4. *Incidental and enjoyable learning via YouTube*

Students in this study observed vocabulary acquisition during emotionally engaging YouTube activities, which underscores the importance of affect in language learning. This supported previous studies in documenting the effectiveness of incidental vocabulary acquisition through audio-visual materials (Hulstijn, 2003; Montero Perez et al., 2014). Peters et al.'s (2016) study is also relevant to the participants' remarked on the ability of enjoyable viewing to facilitate vocabulary remembering and recalling long after the original learning purpose, as a learner described how incidental learning enhances memorization, supporting the idea that there was some affective engagement related to interest and emotion (S3.1.2). It also supports the claim that "affect" is essential for vocabulary retention, which aligns with Krashen's (1982) Affective Filter Hypothesis, which suggests that relaxed environments with strong motivation are conducive to language learners' input uptake. In addition, the development of learners' incidental practices into more systematic structures over time indicates that they can move from casual vocabulary encounters to cyclical intentional learning, or shift along a spectrum from passive to active involvement. Such shifts reflect what AT refers to as emergent object construction, in which learners redefine the object of their activity through sustained interaction with the mediating tool (Engeström, 2015).

In essence, incidental learning on YouTube is not random. It is mediated by learners' interests, habits, emotional responses, and digital environments. The platform's algorithmic personalization and abundance of diverse content types create a dynamic, learner-responsive environment where vocabulary learning becomes a byproduct of meaningful engagement. These conclusions reinforce the argument proposed by Zhang and Zou (2020) regarding informality and YouTube: the spontaneous and enjoyable nature of YouTube allows learners to access vocabulary informally without the pressure of formal instructional contexts, facilitating subconscious uptake and later, more deliberate learning.

3.2.5. Repetition, memory, and learner control in informal contexts

There is a clear gap in students' understanding of the role of repetition in vocabulary retention. Actions such as rewatching instructional videos, using flashcards, and using certain review applications support the idea that vocabulary knowledge is strengthened through repeated encounters (Schmitt, 2008; Webb, 2007). These practices also support the informal application of dual coding theory (Paivio, 1991) and retrieval-based learning (Barcroft, 2007). Students also purposefully rewatch videos and use flashcards in accordance with established vocabulary-learning principles that emphasize repeated encounters with new words for enduring acquisition (Nation, 2013; Webb, 2007).

In addition, learners exercised even greater control by incorporating automated review systems, as exemplified by their use of TFlat to set word-recall reminders. This echoes studies that highlight the value of spaced repetition software (SRS) in mobile-assisted vocabulary learning (Godwin-Jones, 2011). The learners' target goal of vocabulary mastery was not achieved through passive exposure but active engagement with tools and content. Flashcards, rewatching videos, and automated reminders reinforce how learners reconfigure the purpose of digital tools as they work to achieve longstanding cognitive objectives, thus further developing the activity system. These practices also contribute to what Lai and Zheng (2018) term "metacognitive orchestration" in informal learning. Rather than relying on teacher-led repetition, students imposed their own structure on fluid digital environments, adapting the episodic nature of video consumption into systematic vocabulary reinforcement. This confirms that repetition in mobile learning contexts is not only possible but can be purposefully designed, provided that learners possess the strategic awareness and digital literacy to manage their own review cycles.

3.2.6. Influencer-led learning and redefining the learning community

The findings from Section 3.1.6 point to a notable shift in how EFL learners conceptualize community, authority, and support in vocabulary acquisition. Instead of relying solely on classroom teachers or peers, a number of learners in this study turned to YouTube influencers and digital educators as their primary models for vocabulary learning strategies and motivational guidance. These EduTubers, or language influencers, create educational content that

reaches millions and often serve as informal teachers or role models for learners, suggesting that educational YouTube creators have become an influential factor in contemporary language learning, with many students using their videos to complement or even substitute for traditional instruction (Pattier, 2021). This shift from institutional to personal digital models supports Bandura's (2001) theory of social learning, where modeling by relatable, competent others enhances self-efficacy. Influencers, through consistent delivery and practical strategies, serve as community actors in learners' activity systems, expanding the definition of "community" beyond classrooms (Engeström, 2015). The learner's comment about teachers only "telling them to copy vocabulary" (S3.2.2) further supports Reinders and Benson's (2017) view that traditional classroom practices often lack the agency and personalization that digital tools afford.

In sum, these findings challenge the boundaries of conventional learning communities and emphasize the centrality of influencer-based input in students' vocabulary development. YouTube content creators may serve not only as role models but also as affective and strategic supports, embedding social learning in para-social formats. These influencers become crucial actors in the learner's activity system of AT, redefining community, expanding agency, and fostering independent learning behaviors within digitally mediated environments.

3.2.7. Spontaneous engagement and emergent learning pathways

Section 3.1.7's findings highlight how spontaneous and exploratory use of YouTube and similar platforms contributes meaningfully to vocabulary acquisition. Unlike structured or strategic learning, this form of engagement is marked by nonlinear, curiosity-driven behavior, typically triggered by algorithmic recommendations, momentary interests, or habitual mobile browsing. Nevertheless, such unstructured activity can yield incidental learning outcomes, highlighting the flexibility of digital environments in supporting emergent forms of learner agency. Even passive scrolling led to saving videos or repeating exposure, showing learners' capacity to adapt informal input into structured practices. These behaviors, moving from spontaneous viewing to more intentional strategies, reflect the "blurred boundary" between formal and informal learning environments described by Benson (2011) and resonate with research on

incidental learning in digital spaces (Lai & Zheng, 2018), which emphasizes that unplanned exposure can still yield substantial learning benefits, especially when learners develop lightweight mechanisms to consolidate what they encounter. Moreover, they exemplify adaptive agency, in which students respond to contextual learning opportunities by developing undemanding yet effective strategies for vocabulary retention (Lai & Morrison, 2013). From the lens of AT, this process aligns with the concept of emergent object formation (Engeström, 2015). Learners initially enter an activity without a clear learning goal but gradually shift their focus to vocabulary learning as they interact with tools (videos, platforms, apps) that mediate their experiences.

In summary, spontaneous engagement with YouTube represents a valid and often underestimated mode of vocabulary learning. It reflects learners' adaptive agency, their ability to leverage digital affordances in ways that are not necessarily planned but nevertheless purposeful. These behaviors affirm that even within algorithmic, entertainment-driven environments, students remain active meaning-makers. Through ongoing interaction with digital tools, they construct emergent learning paths that bridge informal exposure and self-regulated development, an outcome that is deeply consistent with the tenets of AT.

Discussion of research question 2

The integration of the five themes – strategic and intentional learning, incidental and enjoyable learning, repetition and learner control, influencer-led community support, and spontaneous engagement – provides the most comprehensive understanding EFL students' engagement with YouTube for SDL vocabulary learning when there was more than simple exposure. First, they demonstrated strategic, intentional engagement in creating and sequencing activities and in blending YouTube with dictionaries, flashcards, and AI tools, providing considerable evidence of metacognitive planning and self-regulatory control. Additionally, learners experienced incidental and enjoyable learning by acquiring vocabulary through entertainment-driven viewing, which, together with the affective component, enhanced memory and motivation. The use of repetition and learner control described a further layer of engagement, in which practices such as rewatching videos, using flashcards, and automated reminders provided

systematic control to consolidate vocabulary. In addition to cognitive approaches, social interactions were important. Some learners saw YouTube influencers and educators as mentors and role models, indicating a shift from classroom authority figures to broader digital communities that facilitated motivational and supportive vocabulary learning. Finally, spontaneous engagement illustrated that even unplanned and casually curiosity-driven viewing could translate into deep vocabulary learning as learners integrated incidental input into organized frameworks. In sum, the evidence indicates that students' engagement with YouTube was not a one-dimensional activity, but rather the result of a complex interplay of deliberate strategies, incidental engagement, repetitive practice aimed at consolidation, social learning with influencers and peers, and unprompted pathways provided by the digital context. In this regard, the influence and interplay of strategic, social, and emotional pillars in the appropriation of YouTube for self-directed vocabulary learning is strong.

4. CONCLUSION

4.1. Study contributions

This study examined how Vietnamese EFL university students engage in self-directed receptive vocabulary learning using YouTube as a mobile English learning resource. Grounded in AT, the research illuminated the multifaceted ways learners appropriate YouTube to serve their learning goals. The findings reveal that YouTube functions not only as a source of authentic and engaging input but also as a dynamic mediational tool that students personalize, integrate with other resources, and use both intentionally and incidentally to acquire new vocabulary. Seven key themes emerged from the data: strategic and intentional learning, incidental and enjoyable learning, personalized content access, digital tool integration, repetition and reinforcement, peer and influencer support, and spontaneous exploratory use.

The study contributes to the understanding of MALL by demonstrating how mainstream platforms, though not designed for education, can become powerful tools when strategically appropriated by autonomous learners. In this study, AT also provides a useful lens to capture how agency, tools, and sociocultural conditions interact to shape learning outcomes. Ultimately, the findings underscore that effective vocabulary learning via YouTube is not determined by the platform itself,

but by how learners' structure and adapt their engagement with it.

4.2. Limitations and suggestions for future research

The research has some limitations to be acknowledged. First, the sample was limited to a specific population of Vietnamese university students, which may restrict the generalizability of the findings to other contexts. Second, the data were self-reported through journals, interviews, and focus groups, which introduces potential for recall bias or overestimation of vocabulary learning. Third, the study did not include direct measures of vocabulary acquisition, such as pre- and post-tests, nor did it track actual video engagement data through analytics or screen recordings.

Future research may address these limitations by employing mixed-method designs that combine qualitative insights with quantitative assessments of vocabulary gains. Broader participant samples and technological tracking could provide a more comprehensive picture of learners' behaviors and outcomes. Comparative studies across different institutional, linguistic, or cultural contexts could provide a broader understanding of how learners engage with YouTube for vocabulary learning, and longitudinal studies could track vocabulary acquisition and retention over extended periods to measure the impact of sustained YouTube engagement. Nevertheless, this study offers a detailed and situated account of how YouTube mediates vocabulary learning in informal digital environments, and it contributes valuable insights into the evolving practices of self-directed mobile learning.

4.3. Pedagogical implications

The study offers several pedagogical insights that can inform language instruction and curriculum

design. As part of the community and division of labor within the learners' activity system, teachers should recognize the role of learner-preferred tools such as YouTube and consider incorporating them meaningfully into instruction. This may involve curating topic-specific video content, designing reflective tasks based on video viewing, or integrating YouTube-based vocabulary learning into homework assignments. Rather than competing with learners' digital habits, educators can build upon them to reinforce formal instruction.

Moreover, the study highlights the importance of developing learners' strategic and metacognitive skills. While many students demonstrated autonomy, others benefited from explicit strategies, such as using subtitles effectively, managing attention, or linking video input to vocabulary review through tools like Anki or ChatGPT. Instructors can support this process by modeling these practices and embedding strategy instruction into classroom activities. Finally, the finding that learners rely on YouTube influencers and educators' points to the expanding role of non-traditional learning communities. This suggests a need for educators to acknowledge the broader ecosystem in which language learning occurs and to help students critically evaluate and select credible, effective content.

ACKNOWLEDGMENT

The author would like to express sincere thanks to the participants for their participation, which significantly contributes to the research findings. Gratitude is also extended to the editorial board and anonymous reviewers of the intended journal for their valuable feedback and constructive suggestions.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

REFERENCES

- Anderson, M., & Jiang, J. (2018). *Teens, social media & technology 2018*. Pew Research Center. <https://www.pewresearch.org/internet/2018/05/31/teens-social-media-technology-2018/>
- Bandura, A. (2001). Social cognitive theory: An agentic perspective. *Annual review of psychology*, 52(1), 1-26. <https://psycnet.apa.org/doi/10.1146/annurev.psych.52.1.1>
- Barcroft, J. (2007). Effects of opportunities for word retrieval during second language vocabulary learning. *Language Learning*, 57(1), 35-56. <https://doi.org/10.1111/j.1467-9922.2007.00398.x>
- Benson, P. (2011). *Teaching and researching autonomy in language learning* (2nd ed.). Routledge.
- Benson, P. (2015). Commenting to learn: Evidence of language and intercultural learning in comments on YouTube videos. *Language Learning & Technology*, 19(3), 88-105. <http://dx.doi.org/10.125/44435>
- Bhusaery, R., Chaerul, A., & Kamil, A. B. (2024). Investigating students response of using video based

- learning method through YouTube on English vocabulary learning. *INFOTECH journal*, 10(1), 128-131.
<https://doi.org/10.31949/infotech.v10i1.9567>
- Burston, J. (2015). Twenty years of MALL project implementation: A meta-analysis of learning outcomes. *ReCALL*, 27(1), 4-20.
<https://doi.org/10.1017/S0958344014000159>
- DeStefano, D., & LeFevre, J. A. (2007). Cognitive load in hypertext reading: A review. *Computers in human behavior*, 23(3), 1616-1641.
<https://doi.org/10.1016/j.chb.2005.08.012>
- Dizon, G. (2022). YouTube for Second Language Learning: What Does the Research Tell Us?. *Australian Journal of Applied Linguistics*, 5(1), 19-26. <https://doi.org/10.29140/ajal.v5n1.636>
- Ellis, R. (2008). *The study of second language acquisition* (2nd ed.). Oxford: OUP.
- Engeström, Y. (1999). Activity theory and individual and social transformation. In Y. Engeström, R. Miettinen, & R.-L. Punamäki (Eds.), *Perspectives on activity theory* (pp. 19–38). Cambridge University Press.
<https://doi.org/10.1017/CBO9780511812774.003>
- Engeström, Y. (2015). *Learning by expanding: An activity-theoretical approach to developmental research* (2nd ed.). Cambridge University Press.
<https://doi.org/10.1017/CBO9781139814744>
- Fauziah, F., & Diana, N. (2023). Exploring students' informal digital learning of English (IDLE) and self-regulated language learning from a sociocultural perspective. *Indonesian TESOL Journal*, 5(2), 197-214.
<http://dx.doi.org/10.24256/itj.v5i2.4225>
- Garrison, D. R. (1997). Self-directed learning: Toward a comprehensive model. *Adult Education Quarterly*, 48(1), 18–33.
<https://doi.org/10.1177/074171369704800103>
- Godwin-Jones, R. (2011). Emerging technologies: Mobile apps for language learning. *Language Learning & Technology*, 15(2), 2–11.
<http://dx.doi.org/10.125/44244>
- Hajimaghsoodi, A., & Maftoon, P. (2020). The effect of activity theory-based computer-assisted language learning on EFL learners' writing achievement. *Language Teaching Research Quarterly*, 16, 1-21.
<https://files.eric.ed.gov/fulltext/EJ1270378.pdf>
- Heriyanto, D. (2015). The effectiveness of using YouTube for vocabulary mastery. *ETERNAL (English Teaching Journal)*, 6(1), 1-12.
<http://dx.doi.org/10.26877/eternal.v6i1.2290>
- Hulstijn, J. H. (2003). Incidental and intentional learning. *The handbook of second language acquisition* (pp. 349-381). Blackwell Publishing Ltd.
<https://doi.org/10.1002/9780470756492.ch12>
- Kaboocha, R., & Elyas, T. (2015). The impacts of using YouTube videos on learning vocabulary in Saudi EFL classrooms. In *ICERI2015 Proceedings* (pp. 3525-3531). IATED.
<https://library.iated.org/view/KABOUHA2015IMP>
- Kaboocha, R., & Elyas, T. (2018). The effects of YouTube in multimedia instruction for vocabulary learning: Perceptions of EFL students and teachers. *English language teaching*, 11(2), 72-81.
<http://doi.org/10.5539/elt.v11n2p72>
- Kalyuga, S., Ayres, P., Chandler, P., & Sweller, J. (2003). The expertise reversal effect. *Educational Psychologist*, 38(1), 23–31.
https://doi.org/10.1207/S15326985EP3801_4
- Kelsen, B. (2009). Teaching EFL to the iGeneration: A survey of using YouTube as supplementary material with college EFL students in Taiwan. *Computer-Assisted Language Learning Electronic Journal*, 10(2), 10-27.
<https://callej.org/index.php/journal/article/view/183>
- Knowles, M. S. (1975). *Self-directed learning: A guide for learners and teachers*. Association Press.
- Krashen, S. D. (1982). *Principles and practice in second language acquisition*. Pergamon Press.
- Kukulka-Hulme, A., & Viberg, O. (2018). Mobile collaborative language learning: State of the art. *British Journal of Educational Technology*, 49(2), 207–218. <https://doi.org/10.1111/bjet.12580>
- Lai, C., & Gu, M. (2011). Self-regulated out-of-class language learning with technology. *Computer Assisted Language Learning*, 24(4), 317-335.
- Lai, C., & Morrison, B. (2013). Towards an agenda for learner preparation in technology-enhanced language learning environments. *CALICO Journal*, 30(2), 154–162. <http://dx.doi.org/10.11139/cj.30.2.154-162>
- Lai, C., & Zheng, D. (2018). Self-directed use of mobile devices for language learning beyond the classroom. *ReCALL*, 30(3), 299–318.
<https://doi.org/10.1017/S0958344017000258>
- Montero Perez, M., Van Den Noortgate, W., & Desmet, P. (2014). Captioned video for L2 listening and vocabulary learning: A meta-analysis. *System*, 41, 720–739.
<https://doi.org/10.1016/j.system.2013.07.013>
- Musyaffa, F. B., & Hapsari, C. T. (2025). YouTube as Informal Learning Media to Enhance English Vocabulary: Students' Perception. *Bookchapter English Language Teaching, Literature, and Translation*, 1(1), 285-318.
<https://doi.org/10.15294/eltlt.v1i1.406>
- Nation, I. S. P. (2013). *Learning vocabulary in another language* (2nd ed.). Cambridge University Press.
<https://doi.org/10.1017/CBO9781139858656>
- Oxford, R. L. (2011). *Teaching & researching: language learning strategies* (1st ed.). Routledge.
<https://doi.org/10.4324/9781315838816>
- Paivio, A. (1991). Dual coding theory: Retrospect and current status. *Canadian Journal of Psychology/Revue canadienne de psychologie*, 45(3), 255.

- Park, B., Moreno, R., Seufert, T., & Brünken, R. (2011). Cognitive and affective effects of seductive details in multimedia learning. *Computers in Human Behavior*, 27(1), 291–301. <https://doi.org/10.1016/j.chb.2010.08.007>
- Pattier, D. (2021). The gender gap among EduTubers and the factors significantly influencing it. *Journal of New Approaches in Educational Research*, 10(2), 313–329. <https://doi.org/10.7821/naer.2021.7.732>
- Peters, E., Heynen, E., & Puimège, E. (2016). Learning vocabulary through audiovisual input: The differential effect of L1 subtitles and captions. *System*, 63, 134–148. <https://doi.org/10.1016/j.system.2016.10.002>
- Reinders, H. (2014). Personal learning environments for supporting out-of-class language learning. *English Teaching Forum*, 52(4), 14–19. <https://files.eric.ed.gov/fulltext/EJ1050245.pdf>
- Reinders, H., & Benson, P. (2017). Research agenda: Language learning beyond the classroom. *Language Teaching*, 50(4), 561–578. <https://doi.org/10.1017/S0261444817000192>
- Reinders, H., & White, C. (2016). 20 years of autonomy and technology: How far have we come and where to next? *Language Learning & Technology*, 20(2), 143–154. <http://dx.doi.org/10.125/44466>
- Schmitt, N. (2008). Review article: Instructed second language vocabulary learning. *Language Teaching Research*, 12(3), 329–363. <https://doi.org/10.1177/1362168808089921>
- Sydorenko, T. (2010). *Modality of input and vocabulary acquisition*. <https://scholarspace.manoa.hawaii.edu/server/api/core/bitstreams/ad531755-412d-4a17-8bf7-c722d0e32f13/content>
- Ward, A. F., Duke, K., Gneezy, A., & Bos, M. W. (2017). Brain drain: The mere presence of one's own smartphone reduces available cognitive capacity. *Journal of the Association for Consumer Research*, 2(2), 140–154. <https://doi.org/10.1086/691462>
- Watkins, J., & Wilkins, M. (2011). Using YouTube in the EFL classroom. *Language Education in Asia*, 2(1), 113–119. https://leia.org/LEiA/LEiA%20VOLUMES/Download/LEiA_V2_I1_2011/LEiA_V2I1A09_Watkins_Wilkins.pdf
- Webb, S. (2007). The effects of repetition on vocabulary knowledge. *Applied Linguistics*, 28(1), 46–65. <https://doi.org/10.1093/applin/aml048>
- Zhang, R., & Zou, D. (2020). Types, purposes, and effectiveness of state-of-the-art technologies for second and foreign language learning. *Computer Assisted Language Learning*, 35(1–2), 1–45. <https://doi.org/10.1080/09588221.2020.1744666>