Journal of Ecohumanism
2025
Volume: 4, No: 3, pp. 93 – 116
ISSN: 2752-6798 (Print) | ISSN 2752-6801 (Online)
https://ecohumanism.co.uk/joe/ecohumanism
DOI: https://doi.org/10.62754/joe.v4i3.6317

A Comparison of MALL Listening Input Modes for Immersive Self-Learning Among EFL Learners in China

Xinran Tian[1], Boon Sim Ng[2]*, Hazlina Abdul Halim[3], Jiali Sun[4]

Abstract

This study seeks to explore how differed MALL listening input modes would affect immersive self-learning among English learners in China. A combination of qualitative and quantitative research method is used among 30 first-year university students with a major of English. Through quasi-experimental design and structured interview as well as Mean (x), standard deviation (S.D.), t-test dependable, and content analysis, critical findings are drawn. The study results indicate that video, video and audio as well as video and ASR text input modes could affect incidental vocabulary acquisition and listening comprehension, as moderated by self efficacy and preferred listening style. This research provides insights for EFL students in immersive self-learning.

Keywords: *MALL Immersive Self Learning*、*Listening Comprehension*、*Incidental Vocabulary*、*Listening Input Modes.*

Introduction

The development of mobile technology and applications has shifted the way language is taught and learned in modern times, especially for non-native speakers of English as a foreign language (EFL). In mobile era, mobile assisted language learning (MALL) as supported and facilitated by mobile devices such as mobile phones, tablets, pads, interactive and smart boards, wearable devices and various applications (Ezra & Cohen, 2018) are including adopted by EFL learners. Considering its convenience and perceived benefits, MALL teaching and learning is increasingly utilized among EFL learners (Cerezo, Calderón & Vicente, 2019). Not surprisingly, given safety and hygiene needs post COVID-19 and weighing the costs of commuting to classrooms, we are seeing a noticeable rise in MALL learning among EFL learners (Karakaya & Bozkurt, 2022). Under that backdrop, educators are taking advantage of updated technology to improve students performance in their curriculum design (García Botero, Questier & Zhu, 2019). At the same time, learners are also exploiting various functions of mobile devices in their learning journey (Zou, Li & Li, 2018). More specifically speaking, more learners are using mobile devices and applications to learn in an interactive and dynamic manner by themselves, also known as immersive self learning (Freina & Ott, 2015). Immersive self learning effectiveness has further been examined in aspects of knowledge and skill development. According to empirical research of Freina and Ott (2015), digital content and environment replicating life scenarios could enable learners to become active participants to facilitate outcomes. In light of this, there are increasing research examining effects of MALL on EFL learners who are conducting immersive self learning. Despite that, findings of the previous research are not conclusive in defining how MALL input modes could affect immersive self-learning among EFL learners. Limited research is carried out to research how different input modes could affect learning attitudes, learning processes and learning performances among EFL learners. Research with a focus on listening aspect is also limited, providing gap for the research. To further explore the gap and contribute to future research in this area, this research would compare the impact of different MALL listening input modes on immersive learning outcomes among EFL learners. The research is significant as it may provide critical implications for EFL learners in

[1] Department of Foreign Language, Faculty of Modern Languages and Communication, Universiti Putra Malaysia, 43400, UPM, Serdang, Selangor, Malaysia, Department of Basic Courses, Harbin Institute of Information Technology, Harbin 150431, Heilongjiang Province, China, Email: gs68667@student.upm.edu.my
[2] Department of Foreign Language, Faculty of Modern Languages and Communication, Universiti Putra Malaysia, 43400, UPM, Serdang, Selangor, Malaysia, Email: ngboon@upm.edu.my (Corresponding Author)
[3] Department of Foreign Language, Faculty of Modern Languages and Communication, Universiti Putra Malaysia, 43400, UPM, Serdang, Selangor, Malaysia, Email:hazlina_ah@upm.edu.my.
[4] Department of Basic Courses, Harbin Institute of Information Technology, Harbin 150431, Heilongjiang Province, China, Email: Sunjiali0801@163.com

Journal of Ecohumanism
2025
Volume: 4, No: 3, pp. 93 – 116
ISSN: 2752-6798 (Print) | ISSN 2752-6801 (Online)
https://ecohumanism.co.uk/joe/ecohumanism
DOI: https://doi.org/10.62754/joe.v4i3.6317

China when listening in the self-immersive environment.

Literature Review

MALL Learning Effectiveness Among EFL Learners

The effectiveness of MALL environment on EFL learners has been explored widely. With help of various and updated mobile devices, learners could access learning materials more easily and interactively to support their learning process, as compared to conventional teaching and learning (Ezra & Cohen, 2018). In another research, MALL is claimed to contribute to language acquisition readiness of learners as well, given a dramatic rise in psychological variables (Nuraeni et al., 2020). To be more specific, learners could be more intrinsically motivated and would be more autonomous in the learning process to enhance learning greatly (Nuraeni et al., 2020).

For EFL learners, MALL environment is also found out to be beneficial. It is concluded by Soleimani, Ismail and Mustaffa (2014) that MALL could contribute to extensive reading and vocabulary development of EFL learners when it comes to a foreign language. In another study, MALL effects are observed through reduced anxiety of learners (Rahamat et al., 2017) . Besides, it is also proposed that autonomy and control of learners could be dramatically increased in MALL context (Rahamat et al., 2017).

However, the limitations of MALL learning has also been put forward. It is indicated that that a lack of pedagogical justifications and limited human contact would affect quality of learning (Dawson et al., 2018). It is also claimed that multi-tasking functions of mobile phones could hinder focus and attention to harm learning efficiency (Rajendran & Yunus, 2021).

Listening Input Modes for EFL Learners

As a crucial aspects of learning and a critical way of receiving messages, researchers have examined ways to maximize learning outcomes in listening aspect. According to Huong and Abbott (2017), listening enables learners to internalize core linguistic and cognitive processing skills. However, listening is a challenge for EFL learners as they have to convert sound stream into information units real time (Danan, 2016).

In the listening process, volume, delivery rate, sentence stress, hesitations and pausing, colloquial language and others could all affect cognitive process (Danan, 2016). Other variables identified consist of vocabulary and grammatical understanding, topic familiarity, self efficacy of language learners and listening strategies (Ozcelik et al., 2019; Namaziandost et al., 2019). Among cognitive, meta-cognitive and social affective listening strategies, it is emphasized that meta-cognitive listening strategy could help facilitate listening comprehension and information processing (Ahmadi & Rozati, 2017). Despite task characteristics divergences and individuals differences, directed attention, planning, monitoring and evaluation would be fostered in mega-cognitive process (Ahmadi & Rozati, 2017). The idea is further strengthened by Maftoon and Fakhri Alamdari (2020), claiming that learners would have thorough understanding of the cognitive process and strategy application.

Researchers have also focused on aspects hindering listening processes. Bárkányi (2021) emphasized that language anxiety could be a prominent factor threatening information capture and understanding. Inability to catch with information and speed could threaten comprehension process and listening effectiveness. Despite that, Bárkányi (2021) mentioned that with increased self efficacy, anxiety of EFL learners may be reduced to certain levels to maximize listening outcomes.

In recent years, there are increasing research on listening input modes. As expounded by Lange and Costley (2020), where as visual learners may prefer image and text information processing, auditory learners may better extract information from sounds and speech. Listening comprehension and vocabulary capture would differ among EFL learners when images and transcripts are available or not (Lange & Costley, 2020)

Journal of Ecohumanism
2025
Volume: 4, No: 3, pp. 93 – 116
ISSN: 2752-6798 (Print) | ISSN 2752-6801 (Online)
https://ecohumanism.co.uk/joe/ecohumanism
DOI: https://doi.org/10.62754/joe.v4i3.6317

Self Immersive Learning

As a subjective phenomenological experience, self immersive learning outcomes have been examined. As stated in their article, immersive learning environment could be highly engaging but also cognitive demanding, requiring learners to actively adapt their cognitive, metacognitive and motivational processes (Makransky & Peterson, 2019). In self immersive learning process, self efficacy is considered to be critical to interaction and immersion, as understood as self-held belief and confidence to finish a intended task (Ueki & Takeuchi, 2013).

In their empirical research, Sobocinski, Malmberg and Järvelä (2021) conclude that self immersive learning has critical benefits, with learners fully immersed in the process and regulating themselves in line with goals and objectives. The study of Park, Korbach, and Brünken (2020) generate similar findings, indicating that visual information processing could be enhanced by embodiment, cognitive load and self-regulation.

Research Design

Research Objectives

As previous research has not drawn a conclusive and specific result on the connection between MALL listening input modes and immersive self-learning among EFL learners, the overall objective is to identify the correlation between MALL listening input modes and immersive self learning outcomes among learners. More specifically, the moderating variables that could affect the relation between them would first be pinpointed. On that basis, the research would explore how diverse listening input mode could generate differed outcomes of learning among EFL learners. Finally, specific recommendations would be provided to ESL learners in China in MALL based learning environment to better enhanced their learning outcomes.

Research Framework and Research Hypothesis

Drawing from the above literature, we could conclude that immersive listening outcomes may be affected by individual factors of learner language capacity, preferred listening style and self efficacy. Learner language capacity could affect the linguistic and cognitive process and should be considered in research design. Preferred listening style matters as there may be a divergence between listening input modes and it. Finally, self efficacy affects intrinsic motivation and psychological preparedness of the listeners and should be weighed.

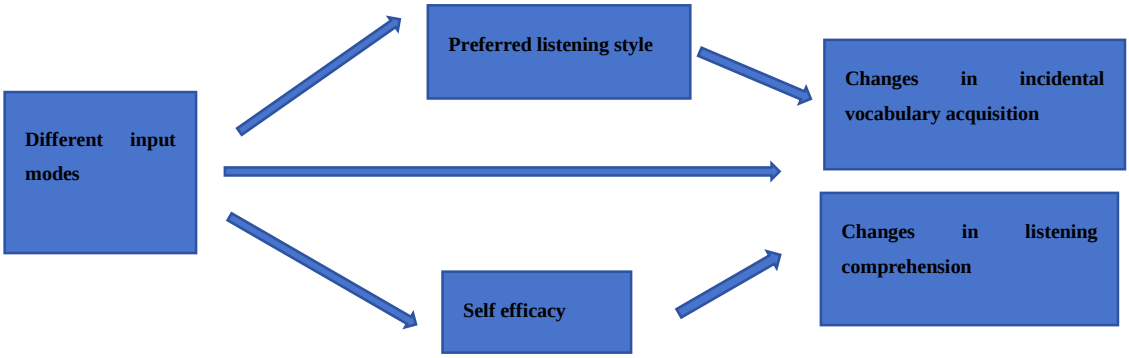
To examine the listening outcomes, listening comprehension and incidental vocabulary acquisition are focused on, as they remain crucial aspects of information processing and learning. In this research, change the above factors before and post the experimental design would be used to measure listening effectiveness.

Besides, self-immersive listening outcomes would also be affected by the listening strategies applied as well. As listening strategy could be difficult to be examined, meta-cognitive strategy could be adopted by all participant.

Finally, video, video and audio and video as well as ASR text input modes would be the three modes considered in research design.

Considering all the preliminary findings and the research objective, the research framework could be crafted as follows.

Journal of Ecohumanism
2025
Volume: 4, No: 3, pp. 93 – 116
ISSN: 2752-6798 (Print) | ISSN 2752-6801 (Online)
https://ecohumanism.co.uk/joe/ecohumanism
DOI: https://doi.org/10.62754/joe.v4i3.6317



As demonstrated in the framework, it is hypothesized that:

Hypothesis 1: Different modes of MALL (Mobile-Assisted Language Learning) listening input will have varying effects on listening outcomes, particularly in terms of incidental vocabulary acquisition and listening comprehension.

Hypothesis 2: The impact of different MALL listening input modes on listening outcomes will be moderated by learners' preferred input mode; learners whose preferred listening style matches the input mode will demonstrate improved listening outcomes.

Hypothesis 3: Different MALL listening input modes will influence learners' self-efficacy, which will, in turn, affect their listening outcomes.

Research Methods

To testify the research hypothesis, a quasi experimental design coupled with a structured interview is used to collect data quantitatively and qualitatively. The experimental component examines the effect of different MALL listening input modes on listening outcomes, while structured interviews provide additional insight into learners' experiences and preferences

Participants

A total of 30 first-year English major students from a 211 university in China are recruited via random sampling. This sample represents a prominent group of EFL learners for whom English is integral to academic success and daily life. Comparing the students' listening outcomes across different listening input modes allows for a dynamic assessment of how these modes impact EFL listening outcomes in China.

Research Instruments

Self-Efficacy Assessment

- Tool: 10-item Likert scale (1–5) self-efficacy questionnaire.
- Purpose: Measures confidence in responding to listening questions based on vocabulary and comprehension of news-related content.

Preferred Listening Style

- Tool: Self-reported survey.

Journal of Ecohumanism
2025
Volume: 4, No: 3, pp. 93 – 116
ISSN: 2752-6798 (Print) | ISSN 2752-6801 (Online)
https://ecohumanism.co.uk/joe/ecohumanism
DOI: https://doi.org/10.62754/joe.v4i3.6317

- Purpose: Participants select a preferred listening input mode (video, video + audio, video + ASR text).

Listening Outcomes Measurement

- Areas: Comprehension and incidental vocabulary acquisition.
- Tests: Comprehension and vocabulary tests developed using content from *China Daily* and *VOA*.
- Design: Tests created by an experienced English teacher and a news editor, rated for difficulty.
- Scoring: Scores on each test range from 0 to 100, based on performance.

Qualitative Insights

- Tool: Structured interview (5 minutes per participant).
- Purpose: Collects personal insights into how different input modes affect MALL-based self-learning.

Experimental Procedure

The study is conducted over 13 days, divided into three experimental phases with distinct listening input modes. Each phase ends with a test, followed by structured interviews.

Pre-Experiment (Day 1)

- All participants complete a baseline listening test (45 minutes) measuring vocabulary and comprehension.

Phase 1: Audio-Only Input (Days 2-4)

- Activity: Participants listen to four 4-5 minute audio segments three times each across the following times:
    - 7:00-8:00 AM
    - 9:30-10:30 AM
    - 3:30-4:30 PM
    - 7:30-8:30 PM
- Strategy: Each audio is listened to:
    - First listen: Focus on the main idea.
    - Second listen: Identify uncertain or unknown words.
    - Third listen: Refine understanding for precision.
- Written Material: Provided after the third listen for reference.
- Test (Day 4): A 45-minute test assesses comprehension and vocabulary.

Journal of Ecohumanism
2025
Volume: 4, No: 3, pp. 93 – 116
ISSN: 2752-6798 (Print) | ISSN 2752-6801 (Online)
https://ecohumanism.co.uk/joe/ecohumanism
DOI: https://doi.org/10.62754/joe.v4i3.6317

Phase 2: Video + Audio Input (Days 5-8)

- Activity: Participants watch and listen to videos with synchronized audio following the same daily schedule and listening strategies as in Phase 1.
- Test (Day 8): Another 45-minute test evaluates comprehension and vocabulary for this input mode.

Phase 3: Video + ASR Text Input (Days 9-12)

- Activity: Participants watch videos with ASR (Automatic Speech Recognition) text displayed, adhering to the same time schedule and listening strategies as in prior phases.
- Test (Day 12): Final 45-minute test measures comprehension and vocabulary outcomes.

Post-Experiment Interviews (Day 13)

- Structured Interviews: Each participant is interviewed for approximately 5 minutes, focusing on their perceptions of how the different input modes influenced their MALL-based learning.

Data Collection & Analysis

Quantitative Analysis

Quantitatively, this paper adopts pre and post-test comparisons, mean and standard deviation calculations, paired sample T-tests and ANOVA tests to analyze to understand how vocabulary acquisition and listening comprehension are shaped by different input modes .

- Pre- and Post-Test Comparisons: For each input mode, pre-test and post-test scores on vocabulary and comprehension are collected.
- Mean and Standard Deviation Calculations: Mean and standard deviation values for each input mode are calculated to measure the average improvement and consistency of outcomes across participants.
- Paired Sample T-Tests: Paired t-tests compare pre- and post-test scores to assess the significance of improvements in vocabulary and comprehension for each input mode.
- ANOVA Tests: ANOVA analysis is conducted to evaluate whether differences in vocabulary acquisition and comprehension scores between input modes are statistically significant.

Besides that, to measure the moderating impact of self-efficacy scores, self-efficacy scores are analyzed as well.

- Self-Efficacy Scale Scores: Participants rate their confidence in tasks related to vocabulary and comprehension on a 5-point Likert scale.
- Correlation Analysis: Correlation between self-efficacy scores and vocabulary/comprehension outcomes for each input mode is conducted to determine if higher self-efficacy aligns with better listening outcomes.

Qualitative Data Analysis

To uncover themes in students' perceptions and experiences with different input modes and how these perceptions relate to listening comprehension and vocabulary acquisition, thematic coding of interview responses is used.

Journal of Ecohumanism
2025
Volume: 4, No: 3, pp. 93 – 116
ISSN: 2752-6798 (Print) | ISSN 2752-6801 (Online)
https://ecohumanism.co.uk/joe/ecohumanism
DOI: https://doi.org/10.62754/joe.v4i3.6317

- Code Development: Responses are initially coded for key concepts, such as *preferred input mode*, *ease of understanding*, *perceived impact on vocabulary*, *impact on comprehension*, and *self-efficacy* influences.

- Thematic Categories: Codes are grouped into themes based on repeated patterns, including alignment with preferred input mode, perceived benefits of input modes and self-efficacy in relation to input mode.

On that basis, frequency and co-occurrence Analysis as well as pattern analysis are used to provide insights to research questions.

- Frequency Analysis: Counts the frequency of themes to identify dominant perceptions across participants.

- Co-Occurrence of Themes: Examines relationships between themes, such as how self-efficacy co-occurs with challenges or success in specific input modes.

- Comparison Across Input Modes: Patterns are identified regarding which modes students found most beneficial for vocabulary versus comprehension.

- Self-Efficacy Influence on Mode Preference: Analysis focuses on whether high or low self-efficacy learners express distinct preferences or challenges with specific input modes.

Results and Analysis

Preferred Listening Style

After the preliminary survey, the preferred listening style of the participants are included as follows.

Preferred listening style	Participants
Audio	2，6，9，16，17，18，21，22
Video and audio	1，5，8，12，13，15，19，23，24，28，29
Video ASR text	3，4，7，10，11，14，20，25，26，27，30

Self-Efficacy Scale

The self-efficacy of participants are recorded in the following table.

Self efficacy item	Number of items	Mean	SD
Understanding the gist	2	3.560	.7395
Understanding the details	2		
Understanding opinions	2		
Working out the meaning of unknown words	2		
Repeating high-frequency and context specific words	2		

Pre-Experiment Scores

Before the quasi-experiment, participants' listening comprehension and incidental vocabulary scores were rated from 0-100. Based on their average scores, nine students scoring below 80, along with the highest scorer, were excluded. Participants 5 and 11, with large discrepancies between comprehension and vocabulary scores, were also not selected. Ultimately, 15 participants continued in the experiment.

Journal of Ecohumanism
2025
Volume: 4, No: 3, pp. 93 – 116
ISSN: 2752-6798 (Print) | ISSN 2752-6801 (Online)
https://ecohumanism.co.uk/joe/ecohumanism
DOI: https://doi.org/10.62754/joe.v4i3.6317

Post Audio-Input Test Scores

After the audio input, the listening comprehension and incidental vocabulary are scored. Further multiply the level of difficulty, the scores are presented in the following table.

Participant	Listening Comprehension	Incidental Vocabulary	Preferred listening style	Self efficacy
Participant 2	82.82	85.85	Audio	2.5
Participant 7	89.89	91.91	Video and ASR text	3.5
Participant 8	84.84	82.82	Video and audio	4.6
Participant 9	95.95	95.95	Audio	4.4
Participant 10	82.82	85.85	Video and ASR text	3.9
Participant 13	90.9	92.92	Video and audio	3.3
Participant 14	85.85	88.88	Video and ASR text	4.2
Participant 15	91.91	92.92	Video and audio	2.7
Participant 18	92.92	89.89	Audio	3.2
Participant 19	87.87	83.83	Video and audio	2.6
Participant 21	92.92	92.92	Audio	4.3
Participant 22	97.97	97.97	Audio	4.5
Participant 25	88.88	89.89	Video and ASR text	3.5
Participant 26	90.9	91.91	Video and ASR text	3.6
Participant 28	88.88	87.87	Video and audio	2.6

Post Video and Audio-Input Test Scores

After the video input, the listening comprehension and incidental vocabulary are scored. Further multiply the level of difficulty, the scores are presented in the following table.

Participant	Listening Comprehension	Incidental Vocabulary	Preferred listening style	Self efficacy
Participant 2	80.185	84.245	Audio	2.5
Participant 7	91.35	92.365	Video and ASR text	3.5
Participant 8	88.305	84.245	Video and audio	4.6
Participant 9	94.395	94.395	Audio	4.4
Participant 10	82.215	85.26	Video and ASR text	3.9
Participant 13	93.38	94.395	Video and audio	3.3
Participant 14	85.26	87.29	Video and ASR text	4.2
Participant 15	92.365	91.35	Video and audio	2.7
Participant 18	90.335	86.275	Audio	3.2
Participant 19	88.305	83.23	Video and audio	2.6
Participant 21	89.32	90.335	Audio	4.3
Participant 22	94.395	95.41	Audio	4.5
Participant 25	88.305	90.335	Video and ASR text	3.5
Participant 26	90.335	91.35	Video and ASR text	3.6

Journal of Ecohumanism
2025
Volume: 4, No: 3, pp. 93 – 116
ISSN: 2752-6798 (Print) | ISSN 2752-6801 (Online)
https://ecohumanism.co.uk/joe/ecohumanism
DOI: https://doi.org/10.62754/joe.v4i3.6317

Participant 28	90.335	88.305	Video and audio	2.6

Post Video and ASR Text-Input Test Scores

The listening outcomes post video and ASR are presented as follows.

Participant	Listening Comprehension	Incidental Vocabulary	Preferred listening style	Self efficacy
Participant 2	79.56	84.66	Audio	2.5
Participant 7	91.8	93.84	Video and ASR text	3.5
Participant 8	86.7	83.64	Video and audio	4.6
Participant 9	91.8	92.82	Audio	4.4
Participant 10	83.64	87.72	Video and ASR text	3.9
Participant 13	90.78	91.8	Video and audio	3.3
Participant 14	87.72	91.8	Video and ASR text	4.2
Participant 15	91.8	91.8	Video and audio	2.7
Participant 18	90.78	87.72	Audio	3.2
Participant 19	87.72	82.62	Video and audio	2.6
Participant 21	88.74	89.76	Audio	4.3
Participant 22	93.84	94.86	Audio	4.5
Participant 25	90.78	91.8	Video and ASR text	3.5
Participant 26	92.82	93.84	Video and ASR text	3.6
Participant 28	88.74	86.7	Video and audio	2.6

Interview Result

A general description of the interview results is as follows:

Reflection on Listening Test Validity

Agreement with Test's Accuracy: 13 participants felt that the listening test accurately reflected their listening comprehension and vocabulary acquisition.

Neutral Responses: 2 participants were unsure if the test reflected their abilities accurately, mentioning that test conditions sometimes vary from real-life listening experiences.

Impact of Self-Efficacy on Listening Outcomes

High Self-Efficacy Learners (4 participants): These learners reported that their self-confidence helped them approach challenging listening inputs positively, enabling them to stay focused during both the training and test stages. They felt prepared to tackle unfamiliar input modes, which they believed enhanced their comprehension and vocabulary.

Moderate Self-Efficacy Learners (7 participants): Learners in this group noted feeling somewhat apprehensive when encountering a change in the input mode, especially when tests were difficult. This nervousness sometimes led to missing details, affecting their outcomes in both training and testing.

Low Self-Efficacy Learners (4 participants): These participants often felt overwhelmed by the listening tasks, particularly when exposed to input modes that didn't align with their preferred style. For example,

Journal of Ecohumanism
2025
Volume: 4, No: 3, pp. 93 – 116
ISSN: 2752-6798 (Print) | ISSN 2752-6801 (Online)
https://ecohumanism.co.uk/joe/ecohumanism
DOI: https://doi.org/10.62754/joe.v4i3.6317

Participant 19 specifically mentioned that switching to audio-only input made cognitive processing challenging, leading to frustration and decreased focus.

Perceptions of Different Input Modes on Comprehension

Preference for Video with ASR Text (6 participants): This group, primarily consisting of those who already preferred video with ASR text as their input mode, emphasized that the combination of visual cues and captions aided their comprehension. They felt that ASR text helped them quickly confirm vocabulary and context, which reduced misunderstandings.

Preference for Video with Audio Only (4 participants): Participants who favored video and audio without captions believed it helped maintain their concentration. They noted that the audio-video format provided sufficient context without the potential distraction of captions.

Preference for Audio-Only (5 participants): All participants who preferred audio-only felt that it encouraged deeper focus and helped them avoid distraction from visual elements. They perceived that pure audio input allowed them to actively engage with listening material and strengthened their auditory processing skills.

Perceptions of Different Input Modes on Vocabulary Acquisition

Preference for Video with ASR Text (8 participants): Among these participants, 5 already preferred video with ASR text, 2 preferred video and audio, and 1 preferred audio. They found that ASR captions made it easier to pick up new vocabulary as they could visually confirm and remember unfamiliar words immediately.

Preference for Video with Audio (3 participants): These participants reported that the familiar combination of video and audio without captions helped them reinforce vocabulary by recognizing words within context. They believed that this mode aligned well with their listening habits.

Preference for Audio-Only (4 participants): Those who preferred audio-only noted that they would often mentally record new words as they heard them, later reviewing and memorizing them. This approach, they felt, allowed for repeated exposure to new vocabulary in a manner they found effective for retention.

Alignment of Listening Input Mode with Preferred Style

Convergence with Preferred Style: 11 participants said that practicing in their preferred listening style helped them feel more comfortable and prepared during both training and testing. They noted that familiar input modes minimized cognitive strain, allowing them to perform at their best.

No Conclusive Opinion: 4 participants expressed no strong preference, with some explaining that they could adapt to different input modes without a significant impact on their confidence.

Divergence from Preferred Listening Style

Barrier to Performance: 9 participants found that divergence from their preferred style created a barrier in both training and testing. They mentioned that adjusting to an unfamiliar mode often took additional effort and impacted their focus.

No Significant Impact: 2 participants (whose preferred style was audio) felt that a shift in input mode didn't hinder their learning or performance. They reported a high level of flexibility in adapting to different modes.

Neutral Opinions: 4 participants were inconclusive, indicating that while different input modes had some impact, they could adapt with minimal difficulty.

Journal of Ecohumanism
2025
Volume: 4, No: 3, pp. 93 – 116
ISSN: 2752-6798 (Print) | ISSN 2752-6801 (Online)
https://ecohumanism.co.uk/joe/ecohumanism
DOI: https://doi.org/10.62754/joe.v4i3.6317

Data Analysis and Findings

Qualitative Analysis

Self-Efficacy as a Critical Factor in Listening Outcomes

The interviews strongly suggest that self-efficacy influences how participants adapt to and perform with different input modes.

High Self-Efficacy: Participants with high self-efficacy displayed resilience and motivation when presented with challenging or unfamiliar input modes. They reported being more focused and adaptable, which positively impacted their performance.

Participant 7, who scored high in self-efficacy, shared: "I felt confident tackling the new input modes, even when it was just audio. I just saw it as a challenge to push myself and focus harder."

Participant 14 echoed this sentiment, stating, "When I feel prepared, it doesn't matter if the input mode changes—I know I can adapt."

Moderate to Low Self-Efficacy: Conversely, participants with lower self-efficacy reported that changes in input modes led to stress and cognitive strain, negatively affecting their comprehension and vocabulary acquisition.

Participant 19, who expressed low self-efficacy, mentioned, "When it switched to audio-only, I found myself getting lost and frustrated. Without visuals, I couldn't keep up as well, and I started doubting myself."

Participant 10, with moderate self-efficacy, commented, "I felt okay with video, but switching to audio was challenging. I became anxious and missed some details, especially when the content got harder."

These responses indicate that self-efficacy acts as a moderating factor, helping those with high self-confidence to navigate new input modes successfully, while lower self-efficacy learners may experience setbacks, especially under pressure.

Impact of Input Mode Convergence and Divergence with Preferred Style

The alignment (convergence) or misalignment (divergence) between participants' preferred listening style and the assigned input mode also affected their performance.

Convergence with Preferred Style: Most participants found that convergence with their preferred style enhanced their listening outcomes, as it minimized cognitive load and allowed them to engage with the material more comfortably.

Participant 22, whose preferred style is audio-only, explained, "When it's just audio, I don't get distracted, and I'm able to listen more intently. It's what I'm used to, so I feel more confident in my answers."

Participant 8, who prefers video with ASR text, noted, "Having captions alongside the video feels natural. I don't have to work as hard to understand the context and vocabulary—it just flows."

Divergence from Preferred Style: When the input mode diverged from their usual style, many participants encountered difficulties, describing the experience as mentally taxing.

Participant 5, whose preferred style is video with audio, said, "Switching to audio-only was a challenge. I felt like I was missing out on the visual cues, and it made me second-guess myself."

Participant 3, who favors ASR, mentioned, "Without captions, I feel lost. I end up focusing too much on

Journal of Ecohumanism
2025
Volume: 4, No: 3, pp. 93 – 116
ISSN: 2752-6798 (Print) | ISSN 2752-6801 (Online)
https://ecohumanism.co.uk/joe/ecohumanism
DOI: https://doi.org/10.62754/joe.v4i3.6317

words I don't know, and it slows me down."

These findings suggest that alignment with preferred input modes may lower cognitive load, while divergence can create additional metacognitive barriers, making it harder for students to process information effectively.

Perceptions of Specific Input Modes on Comprehension and Vocabulary Acquisition

When comparing the three input modes, most participants identified video with ASR text as particularly helpful for vocabulary acquisition, while preferences for comprehension varied. However, individual preferences did play a role in how effective each input mode was perceived to be.

Video with ASR Text for Vocabulary Acquisition: Many participants indicated that the ASR text was invaluable for vocabulary learning, as it allowed for real-time word recognition and reinforcement.

Participant 25 shared, "The captions help me remember new words right away. I can connect the pronunciation to the spelling, which helps it stick."

Participant 13, who prefers audio-only, admitted, "Even though I usually don't like captions, they did help with vocabulary. I could catch words I wouldn't have understood otherwise."

Audio-Only for Enhanced Focus: Those who favored audio-only input reported that the lack of visual distractions allowed them to focus on listening and comprehension.

Participant 21, an audio-preferring learner, stated, "Audio helps me stay locked in on what I'm hearing without any visual distractions. It pushes me to listen more carefully and catch the details."

Participant 18 added, "Without visuals, I'm forced to rely on my auditory skills, which makes me more aware of language nuances."

Video with Audio for Contextual Comprehension: Participants who preferred video with audio often noted that visual cues provided helpful context for understanding content, especially with more complex material.

Participant 9 explained, "Seeing the speaker's expressions and gestures adds context. I'm able to understand more, especially when the topic is complicated."

Participant 26, a visual learner, highlighted, "The combination of video and audio is ideal for comprehension. I feel like I get the whole picture, and it makes it easier to stay engaged."

These insights indicate that while video with ASR text is widely perceived as effective for vocabulary acquisition, individual learning preferences influence the perceived effectiveness of each mode for comprehension.

Qualitative Analysis

Group Statistics-Audio Input

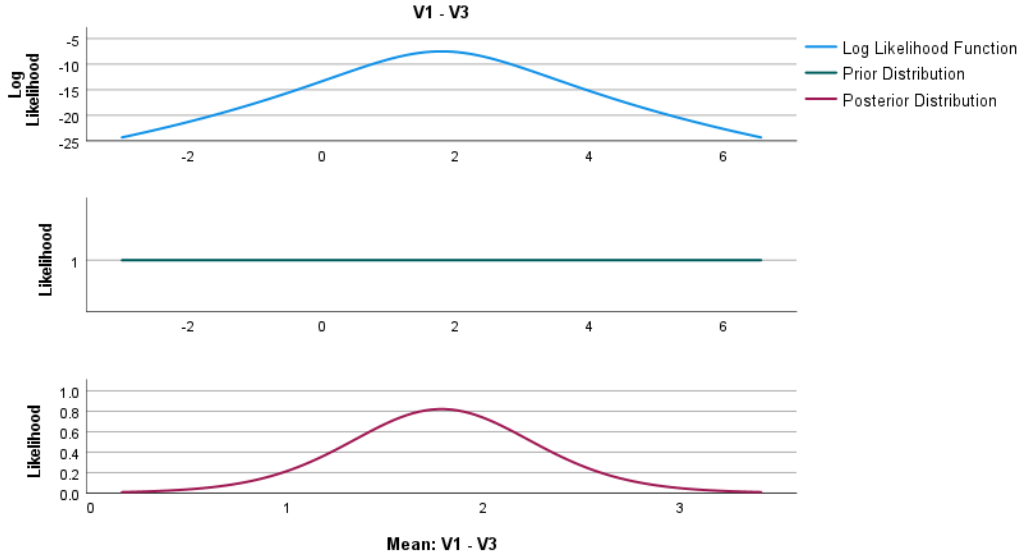
Items	Preferred listening style	N	Mean	Std. Deviation	Std.Error Mean
Pre-test comprehension	1	5	88.00	4.301	1.924
	2	5	88.40	3.362	1.503
	3	5	86.00	4.301	1.924
Post-test comprehension	1	5	92.52	5.828	2.607
	2	5	88.88	2.766	1.237
	3	5	87.67	3.304	1.478
Pre-test vocabulary	1	5	88.80	3.194	1.428

Journal of Ecohumanism
2025
Volume: 4, No: 3, pp. 93 – 116
ISSN: 2752-6798 (Print) | ISSN 2752-6801 (Online)
https://ecohumanism.co.uk/joe/ecohumanism
DOI: https://doi.org/10.62754/joe.v4i3.6317

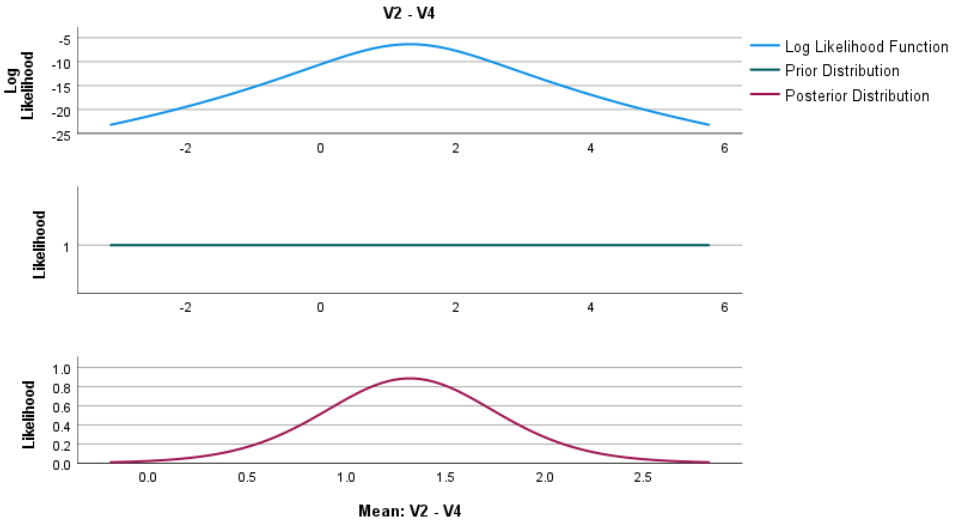
	2	5	86.60	5.128	2.293
	3	5	88.40	2.966	1.327
Post test-vocabulary	1	5	92.52	4.823	2.157
	2	5	88.07	4.812	2.152
	3	5	89.69	2.515	1.125

Paired Group Statistics-Audio Input

		Paired difference					t	df	sig.(2-tailed)
		Mean	Std. Deviation	Std.Error Mean	95% confidence interval for difference lower-bound	upper bound			
Pair 1	Post-test comprehension 1 - Pre-test comprehension	2.221	2.181	.563	1.013	3.429	3.944	14	.001
Pair 2	Post test-vocabulary - Pre-test vocabulary	2.159	1.602	.414	1.272	3.046	5.219	14	.000



Journal of Ecohumanism
2025
Volume: 4, No: 3, pp. 93 – 116
ISSN: 2752-6798 (Print) | ISSN 2752-6801 (Online)
https://ecohumanism.co.uk/joe/ecohumanism
DOI: https://doi.org/10.62754/joe.v4i3.6317



According to mean (x) data above, when the listening input is chosen as audio, the average increase of comprehension and vocabulary is the greatest among people whose preferred style is audio.

Based on paired analysis, when the listening input mode is audio, comprehension before and post the test would change obviously as the p figure is less than 0.001. Besides, the improvement in comprehension is generally more significant than improvement in vocabulary.

Group Statistics-Video with audio Input

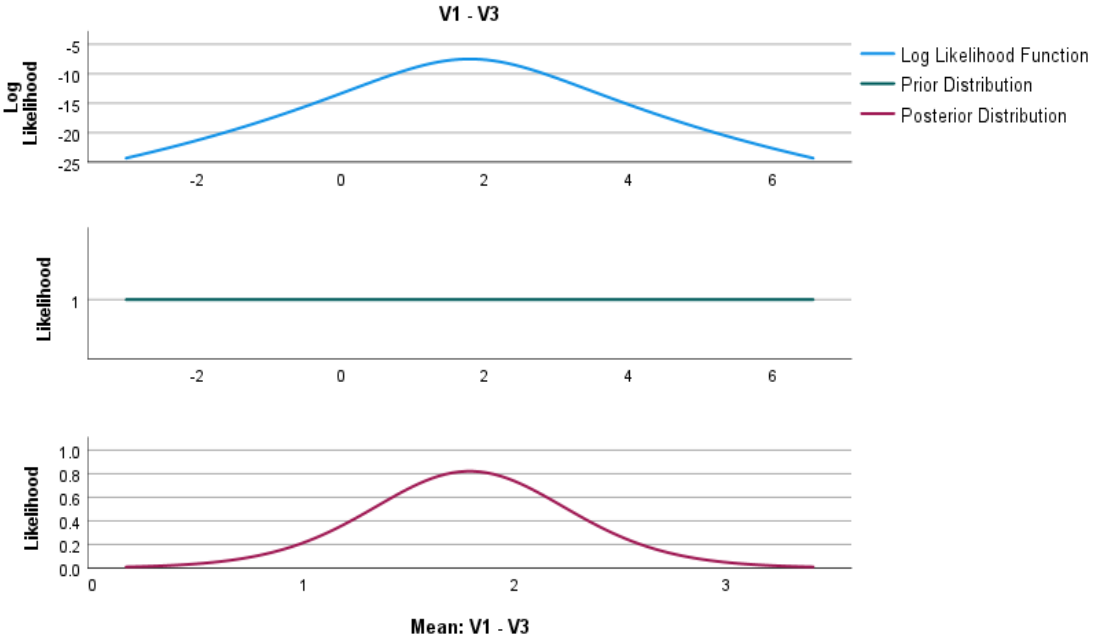
Items	Preferred listening style	N	Mean	Std. Deviation	Std.Error Mean
Pre-test comprehension	1	5	88.00	4.301	1.924
	2	5	88.40	3.362	1.503
	3	5	86.00	4.301	1.924
Post-test comprehension	1	5	89.73	5.813	2.600
	2	5	90.54	2.315	1.035
	3	5	87.49	3.757	1.680
Pre-test vocabulary	1	5	88.80	3.194	1.428
	2	5	86.60	5.128	2.293
	3	5	88.40	2.966	1.327
Post test-vocabulary	1	5	90.13	4.889	2.186
	2	5	88.31	4.706	2.105
	3	5	89.32	2.959	1.323

Paired Group Statistics-Video with audio Input

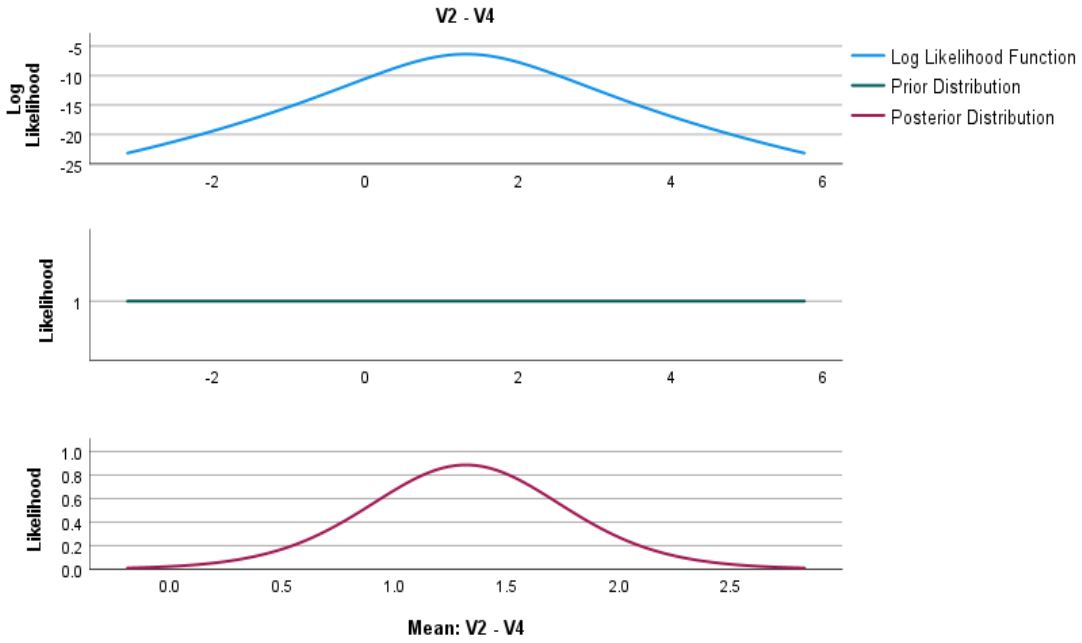
	Paired difference					t	df	sig.(2-tailed)
	Mean	Std. Deviation	Std.Error Mean	95% confidence interval for difference				
				lower-bound	upper bound			

Journal of Ecohumanism
2025
Volume: 4, No: 3, pp. 93 – 116
ISSN: 2752-6798 (Print) | ISSN 2752-6801 (Online)
https://ecohumanism.co.uk/joe/ecohumanism
DOI: https://doi.org/10.62754/joe.v4i3.6317

Pair 1	Post-test comprehension - Pre-test comprehension	1.786	1.705	.440	.841	2.730	4.056	14	.001
Pair 2	Post test-vocabulary - Pre-test vocabulary	1.319	1.580	.408	.444	2.194	3.233	14	.006



Journal of Ecohumanism
2025
Volume: 4, No: 3, pp. 93 – 116
ISSN: 2752-6798 (Print) | ISSN 2752-6801 (Online)
https://ecohumanism.co.uk/joe/ecohumanism
DOI: https://doi.org/10.62754/joe.v4i3.6317



According to mean (x) data above, changing the input mode to video with audio, the average increase of comprehension and vocabulary is the greatest among people whose preferred style is video with audio

Based on paired analysis, when the listening input mode is video with audio, comprehension before and post the test would change greatly as well, as the figure of p is less than 0.006. Besides, the improvement in comprehension is more than that in vocabulary post the training and test.

Group Statistics-Video with ASX Input

Items	Preferred listening style	N	Mean	Std. Deviation	Std.Error Mean
Pre-test comprehension	1	5	88.00	4.301	1.924
	2	5	88.40	3.362	1.503
	3	5	88.00	4.301	1.924
Post-test comprehension	1	5	88.94	5.559	2.486
	2	5	89.15	2.115	.946
	3	5	89.35	3.720	1.664

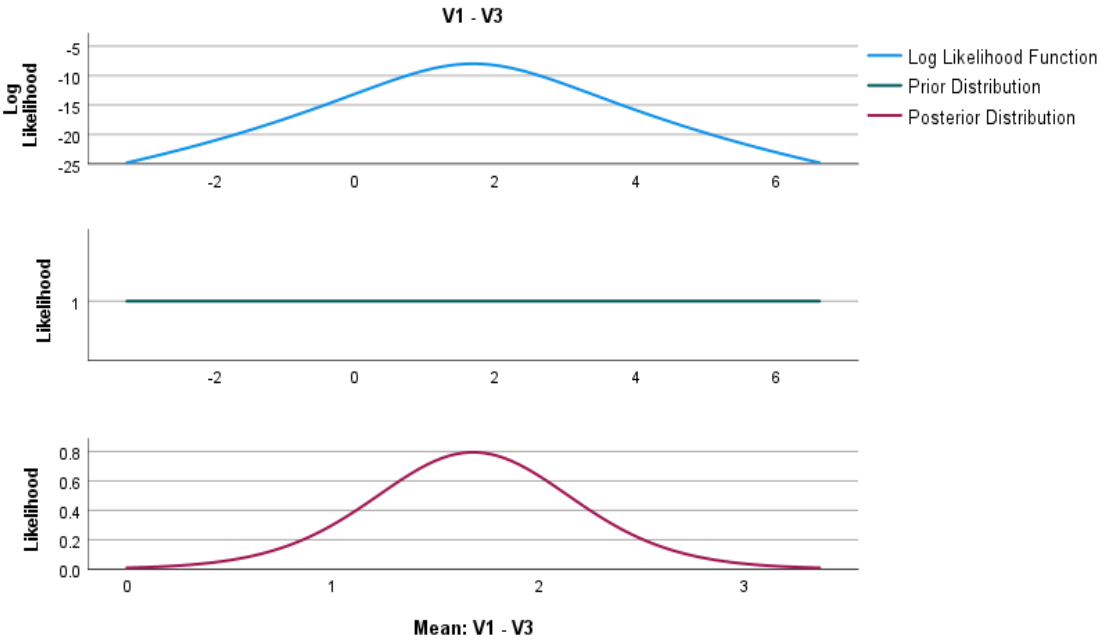
Journal of Ecohumanism
2025
Volume: 4, No: 3, pp. 93 – 116
ISSN: 2752-6798 (Print) | ISSN 2752-6801 (Online)
https://ecohumanism.co.uk/joe/ecohumanism
DOI: https://doi.org/10.62754/joe.v4i3.6317

Pre-test vocabulary	1	5	88.80	3.194	1.428
	2	5	86.60	5.128	2.293
	3	5	88.80	3.194	1.428
Post test-vocabulary	1	5	89.96	4.042	1.807
	2	5	87.31	4.363	1.951
	3	5	89.96	4.042	1.807

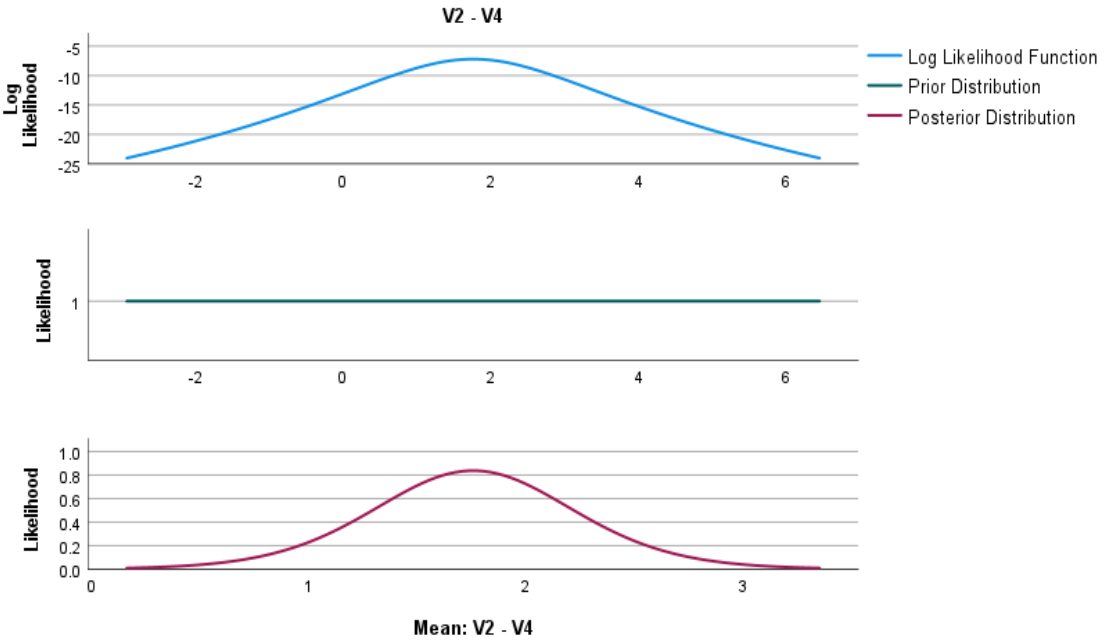
Paired Group Statistics-Video with ASX Input

		Paired difference					t	df	sig.(2-tailed)
		Mean	Std. Deviation	Std.Error Mean	95% confidence interval for difference lower-bound	upper bound			
Pair 1	Post-test comprehension - Pre-test comprehension	1.681	1.762	.455	.706	2.657	3.696	14	.002
Pair 2	Post test-vocabulary - Pre-test vocabulary	1.759	1.672	.432	.833	2.684	4.074	14	.001

Journal of Ecohumanism
2025
Volume: 4, No: 3, pp. 93 – 116
ISSN: 2752-6798 (Print) | ISSN 2752-6801 (Online)
https://ecohumanism.co.uk/joe/ecohumanism
DOI: https://doi.org/10.62754/joe.v4i3.6317



Journal of Ecohumanism
2025
Volume: 4, No: 3, pp. 93 – 116
ISSN: 2752-6798 (Print) | ISSN 2752-6801 (Online)
https://ecohumanism.co.uk/joe/ecohumanism
DOI: https://doi.org/10.62754/joe.v4i3.6317



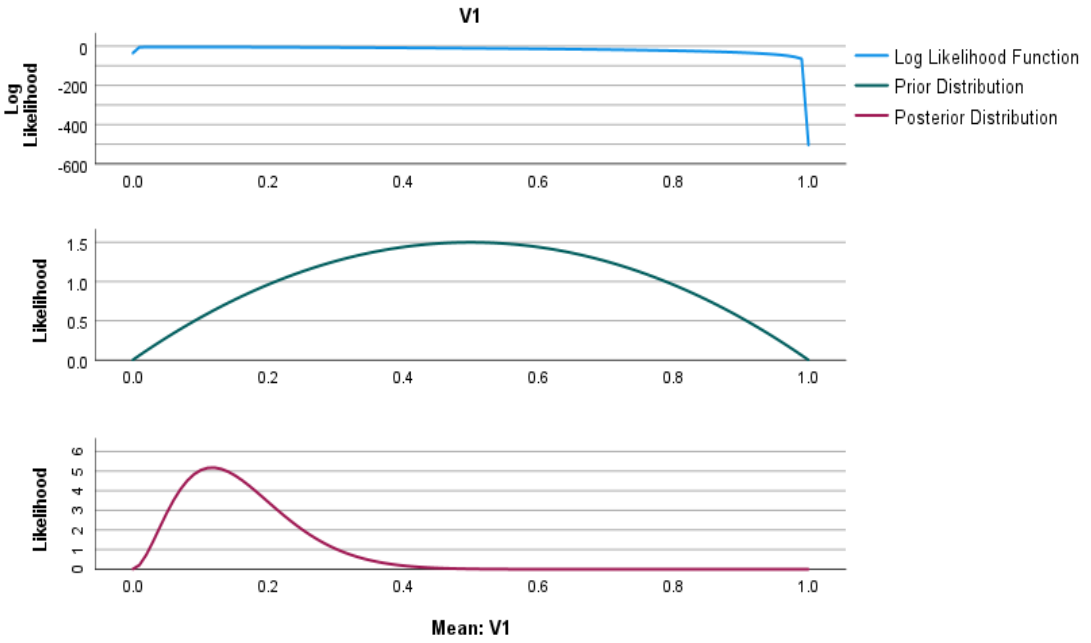
As shown above, when Video with ASX is chosen as input mode, people whose preferred style is also video with ASX would benefit most in listening outcomes, with highest average comprehension and vocabulary increase.

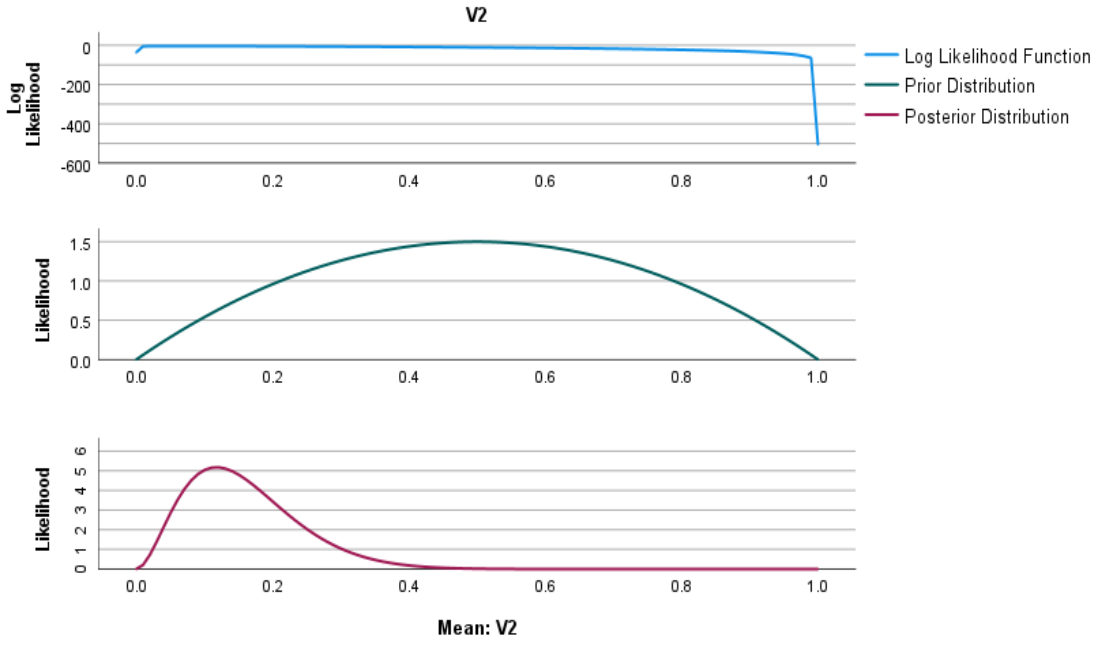
Paired analysis indicate that listening outcomes before and post the video with ASX input differs greatly as the p figure is less than 0.002. Meanwhile, the improvement in vocabulary is more significant than that in comprehension.

Paired Group Statistics-Average post-test Outcomes

		Paired difference					t	df	sig.(2-tailed)
		Mean	Std. Deviation	Std.Error Mean	95% confidence interval for difference				
					lower-bound	upper bound			
Pair 1	Average post-test comprehension-pre test comprehension	1.652604166666691	1.710776515130554	.427694128782638	.740995710234362	2.564212623099021	3.864	15	.002
	Average post-test vocabulary-pre test vocabulary	1.511354166666664	1.502742088557517	.375685522139379	.710599431229305	2.312108902104022	4.023	15	.001

Journal of Ecohumanism
2025
Volume: 4, No: 3, pp. 93 – 116
ISSN: 2752-6798 (Print) | ISSN 2752-6801 (Online)
https://ecohumanism.co.uk/joe/ecohumanism
DOI: https://doi.org/10.62754/joe.v4i3.6317





As shown above, the training of different input modes overall is conducive to listening outcomes in both comprehension and vocabulary.

Journal of Ecohumanism
2025
Volume: 4, No: 3, pp. 93 – 116
ISSN: 2752-6798 (Print) | ISSN 2752-6801 (Online)
https://ecohumanism.co.uk/joe/ecohumanism
DOI: https://doi.org/10.62754/joe.v4i3.6317

ANOVA[a]						
Model		Sum of squares	df	Mean sum of square	F	p
1	regression	28.191	1	28.191	247.727	.000[b]
	Residual error	1.479	13	.114		
	Total	29.670	14			

a. Dependent variable：Average post-test comprehension-pre test comprehension

b. Predictor variable：Self efficacy

Coefficient						
Model		Unstandardized Coefficients		standard Coefficients	t	p
		B	Standard error	Beta		
	V3	1.919	.122	.975	15.739	.000

a. Independent variable：V1

ANOVA[a]						
Model		Sum of squares	df	Mean sum of square	F	p
1	regression	19.319	1	19.319	179.058	.000[b]
	Residual error	1.403	13	.108		
	Total	20.722	14			

a. Dependent variable：Average post-test vocabulary-pre test vocabulary

b. Predictor variable：Self efficacy

Coefficient						
Model		Unstandardized Coefficients		standard Coefficients	t	p
		B	Standard error	Beta		
	V3	1.589	.119	.966	13.381	.000

a. Independent variable：V2

As according to the regression analysis above, it could be concluded that self efficacy could affect the average outcomes in terms of both comprehension and vocabulary post the test, indicating that self efficacy may play a moderating role in the whole process.

Findings

According to the above qualitative and quantitative data analysis, crucial findings could be drawn. Overall, it is implicated that listening input modes could overall enhance listening outcomes in both comprehension and vocabulary acquisition. However, it is found that the convergence between preferred listening style and listening input modes could bring greater benefit to listening outcomes. It is properly explained by the fact that students might have adopted certain listening style for quite long time given preference. Given that, convergence of input mode and listening style is more likely to facilitate meta-cognitive process. It is

Journal of Ecohumanism
2025
Volume: 4, No: 3, pp. 93 – 116
ISSN: 2752-6798 (Print) | ISSN 2752-6801 (Online)
https://ecohumanism.co.uk/joe/ecohumanism
DOI: https://doi.org/10.62754/joe.v4i3.6317

also possible that students having formed certain habit of listening would not expect changes in their listening and studying and would find their cognitive input being changed if the input mode is new to them.

Besides that, it is further find that among the three type of listening input mode, audio as well as audio and video is more conducive to the improvement of comprehension. In the meantime, video with ASX is on average more beneficial to enhancement of vocabulary acquisition, regardless of preferred listening style or self-efficacy. It is probably explained by the fact that ASX would present audiences with the vocabulary, allowing them to acquire unknown words as the listening is processed mega-cognitively.

Finally, it is also found that self-efficacy could play a moderating role. The correlation between high self-efficacy and increased average improvement in listening outcomes of vocabulary and comprehension respectively indicate that internal motivation could facilitate listening outcomes. It is possible that students with higher self-efficacy would be more confident and prepared to receive different listening inputs despite the challenge, thus helping reduce negative factors hindering listening outcomes.

Recommendations

Given the above findings, some implications are drawn for EFL learners in China. First of all, it is suggested that EFL learners in China shall exercise different listening input modes in their daily training, regardless of their preferred listening style. Getting used to different types of listening input could help them get prepared for future training and test as they may be carried out in differed forms. Through that, chances of being intimidated by a new input form could be better reduced to reduce any barriers challenging listening outcomes.

Besides that, it is also recommended that students shall also try to foster self confidence and self-efficacy, as they are conducive to listening outcomes. As a matter of fact, listening to materials with different level of difficulty and finally conquering them as well as being proficient in different learning styles would help them elevate self efficacy.

Finally, for the educators teaching EFL students, crucial implications are also relevant. For them, they should fully consider listening input modes, student efficacy, student preferred styles and other factors in listening instruction, so as to maximize listening outcomes in comprehension and vocabulary acquisition.

Limitations and Implications for Future Research

Despite the relevance of the research in examining relations between input mode and listening outcomes in self-immersive learning environment, there are limitations. To begin with, this research focuses more on a specific group of students, English major students in China. It is difficult to generalize this finding to other groups of students. Besides that, test based approach may not fully reflect the listening outcomes and may hinder the results. Furthermore, given limited time and resources, the sample amount is not big enough, which may hinder the validity of the research to some degree. In the future, it is recommended that how listening input mode would affect listening among different group of students would be analyzed, taking into account age and years of English learning, experiences with listening training, ability to stay concentrated for prolonged period of time as well as other relevant individual and contextual factors.

Conclusion

In conclusion, the research has focused on the impact of listening input mode on EFL students in China who are engaged in immersive self-learning. The research is significant as current research focusing on listening input and outcomes among EFL self-learners is limited. Based on preliminary literature review, the conceptual model is crafted and factors that could affect listening input modes' impact on listening outcomes are identified as preferred listening style and self-efficacy of learners. In quasi-experiment among a group of 15 students, the experiment trained students with differed input modes and recorded their listening outcomes through vocabulary acquisition and listening comprehension. Through Mean (x), standard deviation (S.D.), t-test dependable and regression analysis, it is concluded that students' listening

Journal of Ecohumanism
2025
Volume: 4, No: 3, pp. 93 – 116
ISSN: 2752-6798 (Print) | ISSN 2752-6801 (Online)
https://ecohumanism.co.uk/joe/ecohumanism
DOI: https://doi.org/10.62754/joe.v4i3.6317

outcomes would be affected by input modes and moderated by preferred listening style, self efficacy. Further combining the research result and analysis from the structured interview, it is concluded that the convergence between listening style and input mode as well as higher self-efficacy would lead to better outcomes in comprehension and vocabulary. It is also implicated that whereas audio as well as video with audio method could better contribute to improvement of vocabulary among participants, Video with ASX could better contribute to vocabulary acquisition. On top of the finding, it is implicated that EFL learners shall practice all listening mode in their daily study and should try to enhance self-efficacy through mastering different input modes and getting used to materials with differed difficulty.

References

Ahmadi S.M. & Rozati, F. (2017). The impact of scaffolding and non-scaffolding strategies on the EFL learners listening comprehension development. Journal of Educational Research, 110 (5):447-456. d

Bárkányi Z. (2021). Motivation, self-efficacy beliefs, and speaking anxiety in language MOOCs. ReCALL, 33(2): 143–160.

Cerezo, R., Calderón, V., & Romero, C. (2019). A holographic mobile-based application for practic-ing pronunciation of basic English vocabulary for Spanish speaking children. International Journal of Human-Computer Studies, 124(1):13–25

Danan, M. (2016). Enhancing listening with captions and transcripts. Applied Language Learning, 26, 1-24.

Dawson, P., Henderson, M., Ryan, T., Mahoney, P., Boud, D., Phillips, M., & Molloy, E. (2018). Technology and feedback design. Learning, Design, and Technology: an International Compendium of Theory, Research, Practice, and Policy, 1–45. doi:10.1007/978-3-319-17727-4_124-1

Ezra, O. & Cohen, A. (2018). Contextualised MALL: L2 Chinese students in target and non-target country. Computers & Education, 125: 158-174.

Freina, L.& Ott, M. (2015). A literature review on immersive vir-tual reality in education: State of the art and perspectives, Proc. Int. Sci. Conf. Elearning Softw. Educ., 1(113):1007–1010

García Botero, G., Questier, F., & Zhu, C. (2019). Self-directed language learning in a mobile-assisted, out-of-class context: Do students walk the talk? Computer Assisted Language Learning, 32(1-2), 71–97

Huong, N., & Abbott, M. (2017). Promoting process-oriented listening instruction in the ESL classroom. TESL Canada Journal, 34(1): 72–86.

Karakaya, K., & Bozkurt,A. (2022) Mobile-assisted language learnin

g (MALL) research trends and patterns through bibliometric analysis: Empowering language learners through ubiquitous educational technologies. System, 110, 102925,

Lange C., Costley J. (2020). Improving online video lectures: Learning challenges created by media. International Journal of Educational Technology in Higher Education, 17(1): 1–18.

Maftoon P., Fakhri Alamdari E. (2020). Exploring the effect of metacognitive strategy instruction on metacognitive awareness and listening performance through a process-based approach. International Journal of Listening, 34(1):1–20.

Makransky, G., & Petersen, G. B. (2019). Investigating the process of learning with desktop virtual reality: A structural equation modeling approach. Computers & Education, 134, 15–30.

Makransky, G., Petersen, G.B. (2021). The Cognitive Affective Model of Immersive Learning (CAMIL): a Theoretical Research-Based Model of Learning in Immersive Virtual Reality. Educational Psychology Review, 33: 937–958.

Nuraeni, I.; Carolina, A.; Supriyatna, W.; Widiati, W.; Bahri, S. (2020). Mobile-Assisted Language Learning (MALL): Students' Perception and Problems towards Mobile Learning in English Language. Journal of Physics: Contemporary Series,1641, 012027.

Namaziandost, E., Neisi, L., Mahdavirad, F., Nasri, M., & Monacis, L. (2019). The relationship between listening comprehension problems and strategy usage among advanced EFL learners. Cogent Psychology, 6(1): 1-14. doi: 10.1080/23311908.2019.1691338.

Ozcelik, H. N., Van den Branden, K., & Van Steendam, E. (2019). Listening comprehension problems of FL learners in a peer interactive, self-regulated listening task. International Journal of Listening, 00(00):1–14.

Park, B., Korbach, A., & Brünken, R. (2020). Does thinking-aloud affect learning, visual information process-ing and cognitive load when learning with seductive details as expected from self-regulation perspective? Computers in Human Behavior, 111, 106411

Rajendran, T., & Yunus, M.M. (2021). A Systematic Literature Review on the use of Mobile-assisted Language Learning (MALL) for Enhancing Speaking Skills among ESL and EFL Learners. International Journal of Academic Research in Progressive Education and Development, 10, 586-609.

Rahamat, R., M.Shah, P., Din, R., Abd Aziz, J. (2017). Students' Readiness and Perceptions Towards Using Mobile Technologies For Learning the English Language Literature Component. The English Teacher, 69-84.

Soleimani, E., Ismail, K., & Mustaffa, R. (2014). The acceptance of mobile assisted language learning (MALL) among post graduate ESL students in UKM. Procedia-Social and Behavioral Sciences, 118(1): 457-462.

Sobocinski, M., Malmberg, J., & Järvelä, S. (2021). Exploring adaptation in socially-shared regulation of learn-ing using video and heart rate data. Technology, Knowledge and Learning, 27(2):385–404.

Ueki M., Takeuchi O. (2013). Forming a clearer image of the ideal L2 self: The L2 motivational self system and learner autonomy in a Japanese EFL context. Innovation in Language Learning and Teaching, 7(3): 238–252.

Zou, B., Li, H., & Li, J. (2018). Exploring a curriculum app and a social communication app for EFL learning. Computer Assisted Language Learning, 31(7): 694–713

Journal of Ecohumanism
2025
Volume: 4, No: 3, pp. 93 – 116
ISSN: 2752-6798 (Print) | ISSN 2752-6801 (Online)
https://ecohumanism.co.uk/joe/ecohumanism
DOI: https://doi.org/10.62754/joe.v4i3.6317