Enhancing English Vocabulary Learning: A Comparative Study of Auditory and Visual Input Modalities in ESL

Joe Khoueiry & Rabih Nabhan
joe.h.khoueiry@net.usek.edu.lb   rabihnabhan@usek.edu.lb
Holy Spirit University of Kaslik, Lebanon

Abstract

Vocabulary is a fundamental element in the context of English as a Second Language (ESL). ESL learners find it difficult to develop their language and speaking skills, as well as the comprehension of texts due to limited exposure to the target language outside the classroom, difficulties in pronunciation and intonation, and the struggle to apply vocabulary in different contexts. Despite research efforts to explore different ESL instructional methods, little research was conducted to enhance vocabulary acquisition and retention based on students' learning style and preferences. This study aims to compare the effectiveness of visual and auditory learning methods in improving vocabulary acquisition among ESL students. Using a true-experimental research design, 37 students were selected from the same class and divided randomly into two groups: 20 participants in the Visual-based group and 17 participants in the Auditory-based group. All participants underwent three vocabulary tests: pre-, mid-, and post-tests, and were graded based on performance by two raters. The intervention, conducted over a semester, involved listening to audio clips or watching videos to expand vocabulary. Results indicate significant differences between the groups, with the visual group showing higher vocabulary acquisition scores post-intervention, despite choosing participants from the same class. The Mann-Whitney U test results suggest that the post-test results of the two groups were significantly different, in contrast to the pre- and mid-test score, suggesting that the intervention's effect was obvious in the final test. Friedman's ANOVA conducted for both groups show statistically significant differences within the visual groups across all three tests, unlike within the auditory group. Post-hoc tests reveal that the pre- and post-tests pair in the visual group achieved greater results compared to the other (pre- & mid-, and mid- & post-) pairs. Qualitative data show similar outcomes, proving the power of visual input in enhancing vocabulary. The findings highlight the importance of incorporating visual input modalities in ESL instruction to improve vocabulary learning outcomes, enhance engagement, and vocabulary retention.

Keywords: English as a Second Language (ESL), Vocabulary Acquisition, retention, input modalities, visual, auditory

1. Introduction

English as a Second Language (ESL) learning is a dynamic process that involves various factors influencing language acquisition and proficiency. One aspect of ESL learning that can be considered the foundation of effective communication and use of language, is vocabulary acquisition (Rahmat & Mohandas, 2020). The first step taken by ESL learners is acquiring vocabulary, forming the basis of language competency before getting started with sentence structure or grammar (Alghamdi, 2018). Language

educators face challenges in determining the most effective approaches to help those learners achieve proficiency when it comes to vocabulary acquisition. For ESL learners to feel confident and proficient in using the language, acquiring enough vocabulary is crucial. According to research, knowing approximately 1,000 – 2,000 words guarantees holding a basic conversation, while learning 8,000 words guarantees the ability to hold an advanced one (Harrison, 2023).

Previous research necessitated the use of proven techniques such as context cues and word connections (Kottacheruvu, 2023). Alternative approaches have also been used to enhance vocabulary learning among ESL learners, including Contextual Learning approach. This approach integrates new information into meaningful contexts that can be linked or relevant to students' experiences and real-life situations (Nakata & Elgort, 2020). In addition, using mnemonic devices as a multimodal approach to improve vocabulary retention. This approach was tested in a recent study conducted on learners with Chinese as their native language. These mnemonic techniques, including the keyword and sentence methods, benefit from auditory and visual input to facilitate memory recall, showcasing their pedagogical value in language learning contexts (Hill, 2022). The diverse needs and learning styles of ESL learners are challenges that can get in the way of achieving good results. Limited exposure to the target language outside the classroom, difficulty in pronunciation, and the struggle to use certain vocabulary in the right context, are challenges that demonstrate the importance of adopting effective instructional approaches to get the best outcomes.

1.1. Problem Statement

However, despite the effectiveness of the aforementioned approaches, there is still a gap in understanding how different input modalities, such as auditory and visual-based instructional methods, influence vocabulary acquisition in ESL learners. This gap underlines the need for exploring the specific and different influence of each input modality on vocabulary learning outcomes and speaking improvement among ESL university students.

1.2. Hypothesis Statement

We hypothesize that visual-based learning will lead to greater vocabulary acquisition, memory, and retention compared to audio-based learning among ESL students.

1.3. Purpose of the Study

This study aims to investigate the effectiveness of auditory and visual-based instructional methods in improving vocabulary acquisition and speaking improvement among ESL students. Specifically, the study aims to assess the different impact of these input modalities on vocabulary learning outcomes and speaking skills development.

1.4. Objectives

1. To compare the effectiveness of auditory and visual-based instructional methods in enhancing vocabulary acquisition and speaking improvement among ESL learners.
2. To provide recommendations for educators on the use of input modality methods to boost vocabulary acquisition, retention, and speaking skills in ESL context.

2. Literature Review

The review will focus on ESL learning in general, auditory-based learning approach, and visual-based learning approach. It will emphasize the cognitive and strategic processing involved in vocabulary acquisition and its pedagogical implications.

2.1. Background and Evolution of ESL Learning

With the development of language teaching approaches like the Grammar-Translation Method and the Direct Method in the early 20th century, ESL learning began to emerge (Hilgendorf, 2020). These early methods placed little attention on vocabulary acquisition and conversational skills and focused mostly on grammatical rules.

Language teaching shifted towards communitive strategies when Communicative language Teaching (CLT), where it was first introduced in the late 20th century (Hien, 2021). These methods emphasized the importance of meaningful communication, authentic language use, and interactive learning activities in ESL. Various studies explored the implementation of CLT in classrooms, such as Jabri & Samad (2021), where

CLT was implemented by teachers and linguists at SMPN 1 Enrekang, in Indonesia, through a multifaceted approach. To get the most effective results and guarantee real communication, the teachers followed the core characteristics of CLT, using materials tailored to the student's interests and preferences such as songs and role-plays. They also used computer-based programs, a diverse range of media, listening exercises, in addition to other internet resources, aiming for effective and enjoyable learning. The results were successful in improving the students' language skills and motivating them to learn the language.

Thanks to technology, ESL witnessed a whole change and integration of different multimedia tools, online platforms that encouraged easy learning, and digital resources that are one click away from the student. As a result, new opportunities to provide auditory and visual input in language learning contexts have emerged (*Technology in ESL Classrooms: A Modern Approach - Bay Atlantic University - Washington, D.C.*, 2024).

The benefits of multimodal learning strategies, which use both visual and auditory cues to accommodate a range of learner preferences and improve learning results, are becoming increasingly apparent in modern ESL learning (Ganapathy & Seetharam, 2016). This transition toward multimodal techniques reflects our growing understanding of the cognitive processes underlying language learning and the value of a variety of input modalities for efficient vocabulary development and speaking improvement.

Over the years, researchers have explored different instructional methods to enhance vocabulary acquisition among ESL learners, considering the fact that vocabulary learning was neglected, which lead to issues in language proficiency among ESL learners. Two prominent approaches are auditory-based and visual-based learning methods, each employing distinct modalities to deliver instruction and facilitate learning.

2.2. Auditory-Based Learning

This approach employs sound input such as listening exercises, speech instruction, and other audio materials to enhance listening comprehension skills, pronunciation accuracy, and verbal communication abilities. ESL learners rely heavily on auditory-based learning techniques when trying to be proficient in a language, especially when it comes to vocabulary learning. According to Zhang and Graham (2020), learners

with higher listening proficiency levels tend to acquire significant amount of vocabulary through these methods, including second language (L2) explanations, Codeswitching (CS), and Contrastive focus-on-form (CFoF) approaches. The research focused particularly on the CFoF approach, which includes providing crosslinguistic data about vocabulary items, improving vocabulary acquisition for ESL learners across various competency levels with promising results.

Moreover, another research line has tried several strategic approaches to enhance vocabulary acquisition through time, by highlighting the importance of engagement in listening to get the optimal results from EFL learners (Kacani & Cyfeku, 2015). The approaches included spaced intervals with repetition; instead of overloading them all at once, space repetition involves rewording words or concepts at increasing intervals. For example, in a listening activity, learners may encounter a new word and then, after a few hours, review it again, a day later, a week later, and so on, enhancing memory and retention (Bariuad, 2022). Elaborative Repetition, which delves into the nuances of a word by examining its connotations, collocations, and usage in other settings, and is beyond just mere repetition (Msw, 2022). Structured Repetition Activities, such as the "Hand Computer" technique, in which words are grouped into categories according to how familiar they are, reviewing these sections on a regular basis, and moving words between sections as students gain more vocabulary (Al-Homoud & Osman, 2015). And lastly, Integration with Listening Activities. Using this approach, vocabulary is taught through listening activities that incorporate new words into spoken language. For instance, learners can listen to stories and dialogues that use the vocabulary of the target language. Then, they can answer comprehension questions and tasks that require them to incorporate the vocabulary they learned in their answers (Valentini et al., 2018).

2.3. Visual-Based Learning

Visually based learning methods focus on visual input, such as reading materials, multimedia presentations and visual aids. Through visual stimuli, this approach aims to improve reading comprehension, vocabulary retention, and overall language

understanding. A recent study explored the impact of visual-based learning on vocabulary acquisition. Using statistical analysis, including the independent sample t-test and paired sample t-test, researchers investigated this through choosing 60 students from sub-urban schools to learn 45 target words. Researchers then compared pre and post test results between the experimental and control groups. The experimental group outperformed the control group in terms of mean difference in scores (15.62) and overall improvement score (89% higher), indicating considerable improvement in both groups following the intervention. This demonstrated how Visual Vocabulary may help ESL students, especially those with poor language skills, acquire vocabulary (Mohd Tahir et al., 2020). Another research study tested the impact of visual images integrated into writing tasks on vocabulary acquisition for EFL learners through distance learning. Data for the study were gathered using a mixed-methods approach that included semi-open-ended interviews and the Vocabulary Size Test (VST) Monolingual - Version A. The VST's pre- and post-test findings demonstrated a statistically significant improvement in vocabulary knowledge following the use of images in writing tasks (El Khairat, 2021).

2.4. Comparative Studies

A study investigated the effectiveness of written, audio, and audiovisual input on incidental vocabulary learning among university students learning English as a Foreign Language (EFL) in China. Using a pretest-posttest-delayed posttest design with four experimental groups exposed to different input modes: reading a transcript, listening to a documentary, viewing the documentary, and a control group, the study found that significant vocabulary learning occurred through all three input modes, with retention of learned vocabulary one week after exposure. However, there were no significant differences in vocabulary gain and retention between the modes, suggesting that each mode contributed equally to vocabulary acquisition. These findings contribute to understanding the role of diverse input modalities in language learning and support the use of L2 television programs as valuable resources for vocabulary development in EFL contexts (Feng & Webb, 2019).

Several factors could have contributed to the lack of differences in vocabulary acquisition and retention, including having similar cognitive processes involved in

encoding vocabulary from the written, audio, and audiovisual inputs. For example, having the same context clues, exposure time, and repetition, consistent across the different modes. Other factors could have been the sample size, having an effective instructional design, or even testing students with similar proficiency levels, leading to minimized differences in vocabulary learning outcomes across all modes.

2.5. Cognitive and Strategic Processing

Students face several challenges in acquiring vocabulary, using it, and getting to the proficiency level, due to the complexities of language structure and cultural nuances. Therefore, understanding the cognitive processes involved in vocabulary acquisition to get effective outcomes and interventions when it comes to teaching the language and the strategies employed.

In ESL settings, it is essential that learners acknowledge the fact that vocabulary acquisition is not an easy process, but rather a multifaceted one, that seeks to cover the linguistic, cognitive, and socio-cultural aspects (Afzal, 2019). Thus, learners must understand the meaning of the word, the usage context, and syntactic roles within sentences and not just learn it. This is when cognitive processes come into play and shape how vocabulary is acquired, stored, and retrieved in memory. Focusing on cognitive processing, emphasizes the role of mental frameworks in incorporating and arranging new vocabulary within the current linguistic knowledge available.

The process of engaging in operations without conscious awareness automatically, when interpreting auditory text is referred to as Cognitive processing. Processes such as Sematic Encoding, when learners link a word with a mental image or an existing notion, Associative Memory Networks, when learners link a word with a somewhat related concept or term or a contextual cue, aiding in retrieval and retention. Furthermore, Strategic Learning approaches aid in augmenting the cognitive processing, through developing certain techniques such as mnemonic devices, pattern recognition, as well as contextual inference (Saad et al., 2022). Through incorporating these processes into the ESL pedagogy, teachers can further develop effective techniques to facilitate vocabulary acquisition.

3. Methodology

To compare the efficacy of auditory and visual input modalities in improving ESL student's vocabulary learning, this section covers the research design, the participants, the tools used for treatment and intervention, as well as the data collection and analysis procedures.

3.1. Research Design

This study was carried out at University of Balamand in Lebanon, to investigate the impact of Auditory and Visual input modalities on enhancing vocabulary learning, retention, speaking, and memory among university students of *Intensive English II* course. Adding to the pedagogical practices of ESL educators, this study will compare the two input modalities, aiming to find new teaching methods for language and vocabulary acquisition. Additionally, it aims to discover students' preferences and what motivates them to get optimal outcomes.

3.1.1. Participants

A random selection process and a true-experimental research design was used to minimize the differences between the selected groups (Chaudhari, 2021). We chose 37 University students, aged 17-19. Of the sample 35.14% consisted of females, accounting to 13, and 64.9% of them were males, accounting to 24 of the sample. Then, we randomly divided the participants into two groups based on the two-group pre-test post-test design approach.

3.1.2. Course Description

The *Intensive English II* course is an intensive eight-credit course taken by students at different university levels. For the purpose of this research, the intervention was implemented over the course of an entire semester, totaling 90 hours spread across approximately four months. There are no prerequisites required to enroll in this course; students take this course to improve their listening and speaking skills. The students came from a plethora of different majors including engineering, biology, computer science, and many others. They all had similar levels of English proficiency.

3.2. Study Procedure & Data Collection

For the data collection process, students were presented with listening prompts. Subsequently, they were tasked with summarizing the content they had heard. Finally,

they were requested to engage in spoken discourse for approximately two to four minutes on a topic of similar nature. After that, they were given scores based on their performance. Ted Talks were mostly used for the visual Mid and Post-tests. Short 2-4 minutes long audio clips, mostly from a book called *Q: skills for Success* by Colin Ward & Margot Gramer, were used for the Mid and Post-tests for the auditory class.

A. *Pre-Test*

We conducted the same pre-test for all 37 students where no sort of intervention was made, and students were taught speaking and listening skills using the conventional methods. They were sometimes asked to create a PowerPoint presentation on specific topics and present it in class in addition to activities of similar nature. For the pre-test, a selection of comprehension questions was presented to students in different formats, such as oral responses and multiple choice and were graded accordingly.

B. *Mid-Test*

The class was divided into two groups, assigning 17 students randomly to the auditory group and 20 students to the visual group in two different classes. To begin the intervention process, both classes regularly watched videos or listened to audio clips starting from the beginning of the second month until the end of the third. Approximately two months after, each group underwent separate testing sessions. During these tests, students were given 2-4 minutes to either speak on a given topic or summarize an audio clip/video. To ensure that students were given an objective and reliable score, two teachers were present to grade them.

C. *Post-Test*

The third and final test was conducted towards the end of the semester after the intervention had been completed. By that time, the students had already listened to or watched around 30 to 40 audio and video clips covering a wide range of topics to expand their vocabulary as much as possible. Then the students were tested using the exact same method as the mid-test and graded by the same two teachers.

3.3. Data Analysis

After completing the data collection process, a quantitative-qualitative mixed approach was used to analyze the scores of the three tests.

3.3.1. Quantitative Analysis

Before analyzing the data gathered, we needed to make sure that the data met two prerequisite tests: normality and homogeneity. A normality test is done to see if the dataset is normally distributed in the two groups or not as many statistical analyses assume a normal distribution of data such as t-tests and One-Way ANOVA. Using SPSS, we first conducted the Shapiro-Wilk Normality test with a significance criterion (*P-value*) of 0.05 to check the normality of the distribution.

Table 1

Shapiro-Wilk Normality Tests of Audio & Visual Input Modality

Test	Group	p Sig.	Analysis
Pre-test	Audio Input Modality	.025	
	Visual Input Modality	.048	
Mid-Test	Audio Input Modality	.262	Sig. (2-tailed) *value < 0.05.*
	Visual Input Modality	.003	(Mixed distribution)
Post-Test	Audio Input Modality	.512	
	Visual Input Modality	.006	

According to Shapiro-Wilk normality test, a *p* value greater than 0.05 indicates a violation of normality distribution. Table 1 above showed that, in the pre-test, the audio and visual modality groups scored low p-values of .025 and .048 respectively. This indicates that the pre-test data did not follow a normal distribution (p Sig < 0.05). In the mid-test, the audio group scored a high score of 0.262 (*p Sig* > 0.05), while the visual group gained a low score of 0.003 (*p Sig* < 0.05). Moreover, in the post-test, the audio group scored a value of 0.512 (*p Sig* > 0.05), while the visual one scored a value of 0.006 (*p Sig* < 0.05). This shows that in both the mid and post-tests, the data showed a mixed distribution.

Furthermore, to proceed with the analysis, it is important to conduct a homogeneity test to draw accurate conclusions. The homogeneity test, also known as homoscedasticity, assumes that if the variances are homogenous (based on mean >

0.05), the variability within groups is approximately the same across all levels of the independent variable(s).

Table 2

Levene's Test of Equality of Variances for the Audio & Visual Groups

Test	Group	Based on Mean	Analysis
Pre-test	Audio Input Modality	0.426	
	Visual Input Modality		
Mid-Test	Audio Input Modality	0.103	Based on mean > *0.05*
	Visual Input Modality		(Homogenous)
Post-Test	Audio Input Modality	0.397	
	Visual Input Modality		

Table 2 shows the Levene's test of equality of variances' results for both groups. In the pre-test, the homogeneity test obtained a value of 0.426, while the mid-and post-tests scored a value of 0.103 and 0.397, respectively. These results show that the data from the three tests for both groups was homogenous (Tian et al., 2020).

A mixed distribution of normality was exhibited in the data, we therefore concluded that the data was not normally distributed and thus rejected the null hypothesis. Since the normality test is a prerequisite, conducting any parametric analysis was not feasible. Therefore, we opted for non-parametric analysis to ensure accurate results for two of our research questions. Firstly, we used the Friedman's One-Way ANOVA test to compare the pre-test, mid-test, and post-test. This analysis aims to determine if there were significant differences in vocabulary learning performance across the three time points within each group (Mouritsen, 2016). Secondly, we preformed three Mann-Whitney U tests to compare the following pairs of scores within each group:

a) *Pre-test scores in scores in the auditory group versus pre-test scores in the visual group.*

b) *Mid-test scores in scores in the auditory group versus Mid-test scores in the visual group.*

c) *Post-test scores in scores in the auditory group versus Post-test scores in the visual group.*

These tests were conducted to evaluate the difference in vocabulary learning performance at each time point between the two groups. This assessment is crucial for evaluating the effects of intervention in educational research to provide evidence-based decision (Akpan et al., 2023).

3.3.2. Qualitative Analysis

To add more depth to our research, we included qualitative data by gathering feedback from four students of each group. Aiming to maintain consistency and facilitate identification of recurrent themes, we provided both groups with the same set of eight questions. Employing thematic analysis and open coding using SPSS, we aimed to explore students' preferences, perceptions, and experiences with each modality. Our goal was to improve four aspects of vocabulary acquisition and overall learning experiences among ESL learners: engagement, retention, memory, and ease of understanding.

Due to the small sample size and number of questions, we chose consensus coding as a method for assessing inter-rater reliability (Olson et al., 2016). Two independent coders were assigned to code half of the data separately, then asked to discuss their thoughts and reach a discussion on the final coding scheme. As a result, a robust coding framework was developed and applied to the entire sample to ensure reliability and validity. Thus, we coded the qualitative feedback data of the six open-ended questions based on recurrent themes, numbered from 1 to 5 and the two Likert scale questions, numbered from 1 to 5; 1 being Strongly Agree and 5 being Strongly disagree. We then measured frequencies for all questions using descriptive statistics and presented them in bar charts in the result section (Refer to Appendix for the codebook).

4. Results

This section will cover the results of the quantitative data, including the results of Friedman's ANOVA, Bonferroni's Pairwise Comparisons, and the Mann-Whitney U Test. The qualitative data will cover visual bar charts depicting the responses to the eight questions from our participants, highlighting the most prevalent themes in each group.

4.1. Quantitative Data

Our quantitative data aimed to compare the Audio Input Modality group (AIM) and the Visual Input Modality group (VIM), determining if there is a significant difference between the groups in each time period. Table 3 shows the results of Friedman's ANOVA test of the AIM group in comparison to the VIM group.

Table 3

Friedman's ANOVA Results

Group	p-value	Significant
Audio Input Modality	0.055	No
Visual Input Modality	<.001	Yes

The probability values of both groups show that there were no significant differences between the results in pre, mid, and post results in the Audio group, given its probability value of 0.055 that exceeds the conventional threshold of .05. This suggests that there is insufficient evidence to reject the null hypothesis. On the other hand, the visual group scored a value (<.001), resulting in statistically significant differences between the results of the three tests, suggesting that there is sufficient evidence to reject the null hypothesis.

Following the Freidman's test, we conducted the Bonferroni's pairwise comparisons as post-hoc tests for both groups and received a value greater than the conventional threshold of .017 in the audio group, suggesting that no significant difference was found between the three paired comparison. On the other hand, we obtained a value less than the threshold for the three pairs in the visual group, suggesting that change was obvious in three time periods across all three tests conducted.

Table 4

Bonferroni's Pairwise Comparisons as Post-Hoc Tests for Both of the Groups

Group	Pairs	Mean difference	p-value
Audio Input Modality	(Pre-Test, Mid-Test)	0.735	0.330
	(Pre-Test, Post-Test	1.03	0.139

	(Mid-Test, Post-Test)	0.295	0.529
Visual Input Modality	(Pre-Test, Mid-Test)	2.40	<.001
	(Pre-Test, Post-Test	3.525	<.001
	(Mid-Test, Post-Test)	1.125	<.001

Table 4 shows a significant difference in the results of the VIM group. In contrast to the AIM group, which showed no significant differences in results between the pairs of tests, as indicated by the probability value. Moreover, the mean difference in the post-hoc results indicates the average difference in scores between the two groups at that specific time point (pre, mid, & post-tests), providing more evidence and insight into the average magnitude of the differences in scores between the visual and audio groups following the intervention. The Table above shows that the highest mean difference score is in the visual group and between the pre-test and post-test with a score of 3.525 and a p-value of <.001. Whereas in the Audio group, the pre-test & post-test pair scored a value of 1.03 and a p-value of 0.139. This indicates that the visual group scored higher results on the post intervention test, providing us with sufficient evidence to ensure that using visual materials for vocabulary acquisition is more effective than using Audio material. The same thing can be seen when comparing the other two pairs in both groups, indicating that visual material was effective throughout the whole period of treatment.

To verify the reliability of our Friedman's ANOVA test results, we assessed if there was a difference between the test results within each group by comparing the pairs of scores. This involved comparing the distribution of a continuous outcome variable between two independent groups. As the assumptions of a t-test were not met, we conducted the Mann-Whitney U test instead.

Table 5

Mann-Whitney U Test Results for Visual vs. Audio Groups at Each Time Point

Time Point	Mann-Whitney U test	P-value	Interpretation
Pre-Test	.133[b]	> 0.05	No significant difference

Mid-Test	.080[b]	> 0.05	No significant difference
Post-Test	.001[b]	< 0.05	Significant difference

Significance level (α = 0.05)

Table 5 shows the test results of Mann-Whitney U test for the comparison between Visual and Audio groups at each time point (pre, mid, & post-tests). The value of both the pre (.133) and mid (.080) tests are greater than the typical significant level (*p = 0.05*), suggesting that there are no statistically significant differences in the scores of both tests. However, the post-test (.001) p-value was less than the typical 0.05 value, suggesting that the post-test results of the two groups were significantly different. This adds to our results by providing more evidence that the visual group did indeed score more than the audio group after the intervention.

4.2. Qualitative data

Delving deeper into the ESL student's performances, experiences, and preferences, our qualitative data results were in line with our quantitative data, adding more richness to them. Thematic analysis of the responses to the six open-ended questions revealed five distinct patterns each. This section displays visual bar charts depicting the responses to the eight questions from our eight participants, highlighting the most prevalent themes in each group.

1. Open–Ended Questions

Figure 1

Experience with Instructional Method

1.Can you describe your experience with the audio/visual-based instructional method used in this study?
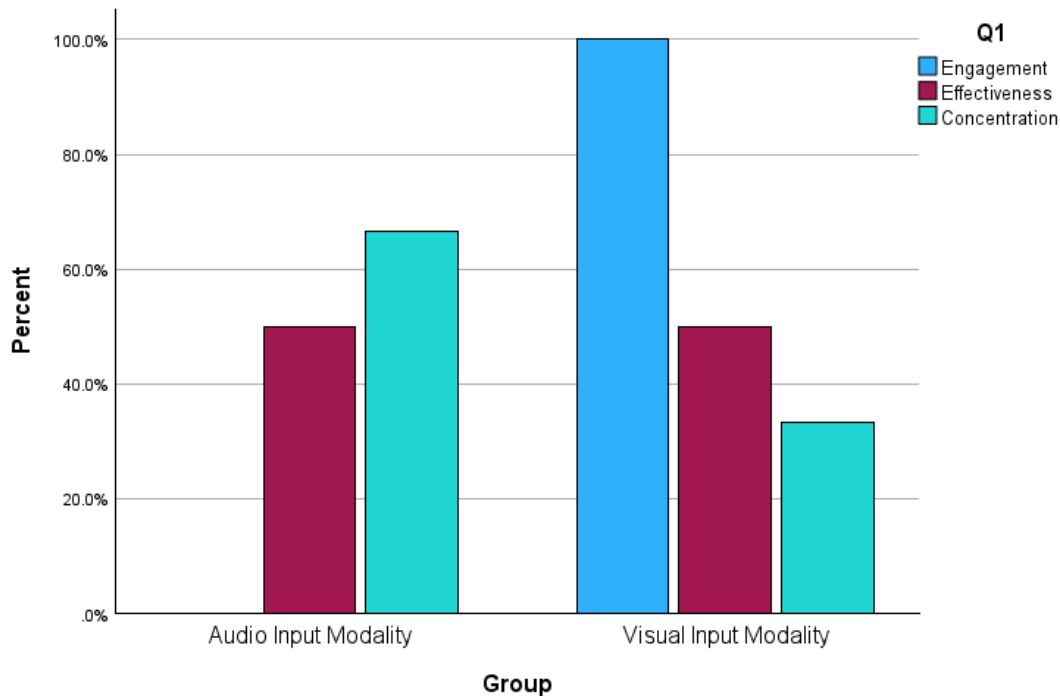
As can be seen in the graphs, students consistently expressed a preference for the visual approach over the auditory one. Figure 1 addresses the experience with the visual approach compared to the auditory approach. Firstly, the issue of engagement stood out. The fact that all participants in the visual group emphasized engagement, while 0% of the audio group made any mention of it, shows the preference for the visual approach due to its perceived higher engagement factor. Surprisingly, the audio group showed a greater emphasis on concentration, with 65% of participants mentioning it, compared to only 30% in the visual group. Lastly, the effectiveness variable is shown in 45% of each group. This indicates an agreement on the importance of this aspect in both visual and auditory learning experiences.

Figure 2

Effectiveness of Instructional Method

2.What aspects of audio/visual-based instruction do you find most effective for learning vocabulary and improving speaking skills?
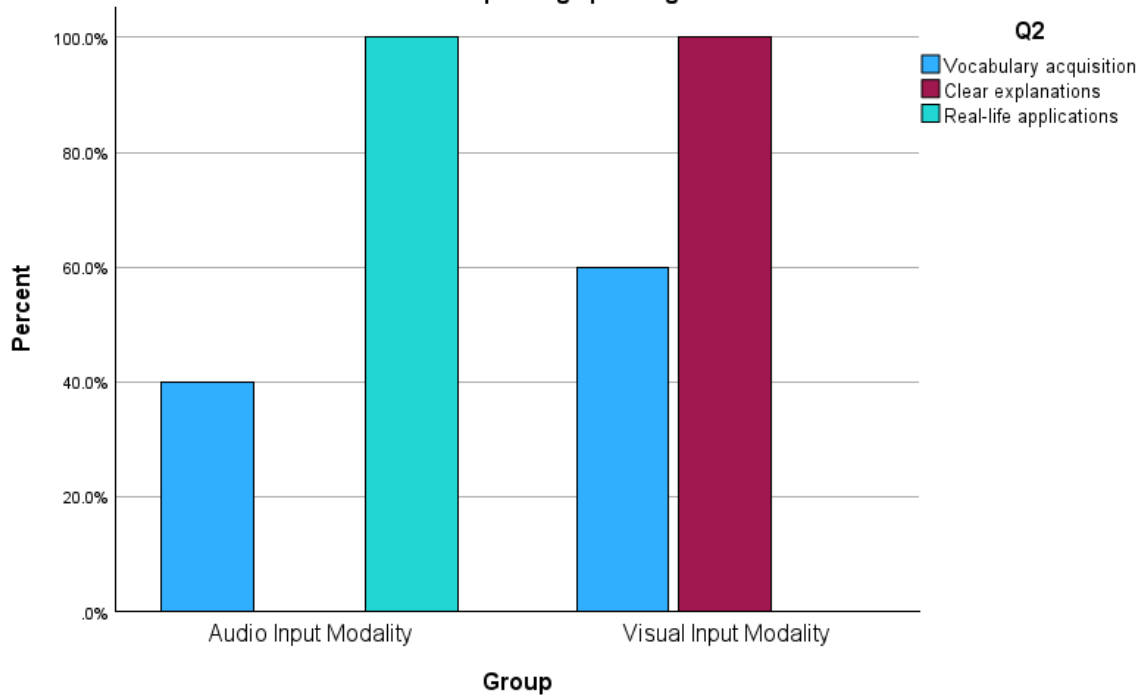

Figure 2 shows the three recurring themes predominant in terms of improving speaking skills: vocabulary acquisition (V 60%, A 40%), clear explanation (V 100%, A 0%), and real-life experience benefit (V 0%, A 100%).

Figure 3

Challenges and Difficulties

3.Have you encountered any challenges or difficulties with the audio/visual-based instructional method? If so, please explain.
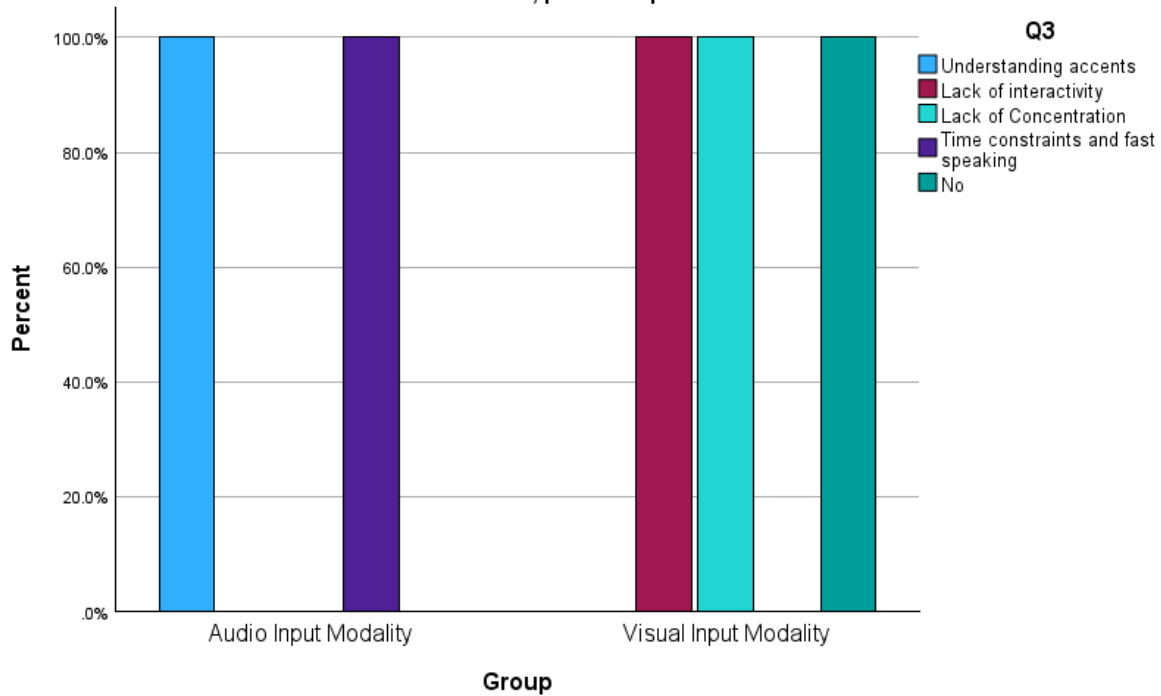
Figure 3 addresses the challenges faced during the sessions. The visual group faced two issues: lack of concentration (100%) and lack of interactivity (100%), while the auditory group faced two others: understanding accents (100%) and coping with time constraints and fast speaking (100%). However, students in V-group noted that they initially experienced these challenges but later adapted, leading to the acquisition of speaking skills and the development of vocabulary, unlike A-group that reported ongoing difficulties with these issues.

Figure 4

Impact on Learning Process

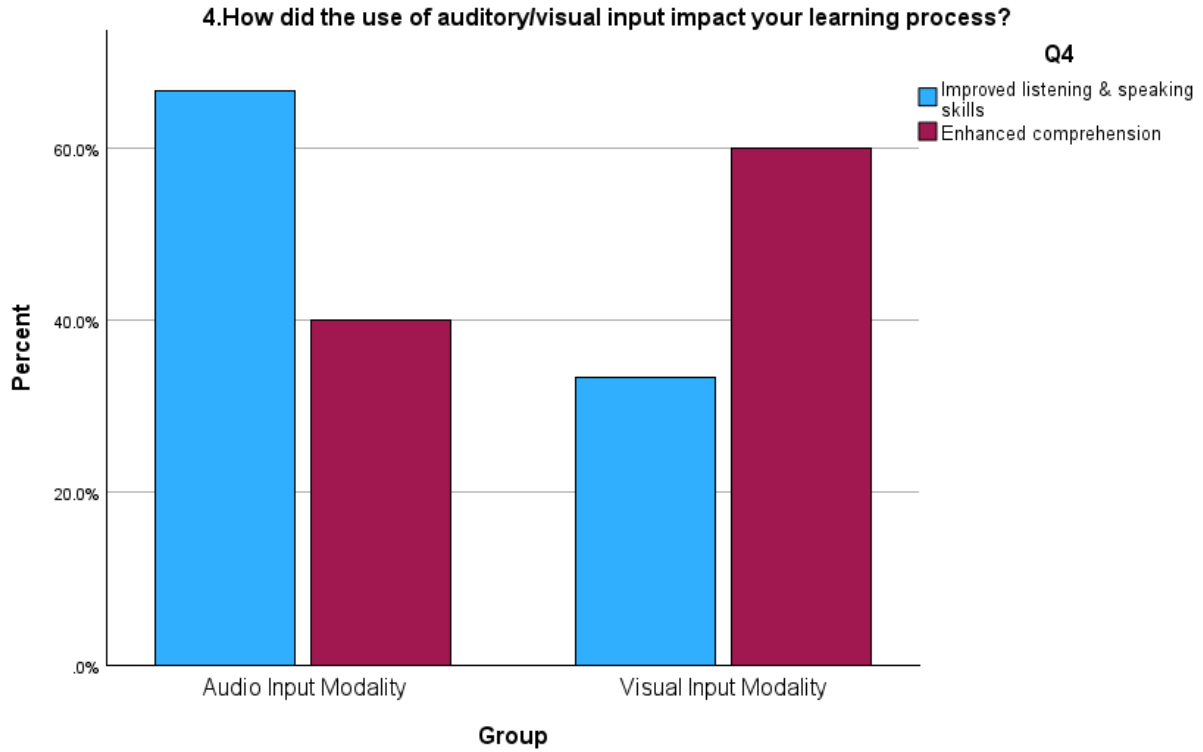4.How did the use of auditory/visual input impact your learning process?

Figure 4 addressed the impact of the learning process, where both groups highlighted: enhanced comprehension (V 60% & A 40%) and improved listening and speaking skills (V 30%, A 65%).

Figure 5

Overall Assessment of Instructional Method

5. Reflecting on your overall experience, do you think the audio/visual-based instructional method was suitable for vocabulary acquisition and speaking improvement? Why?
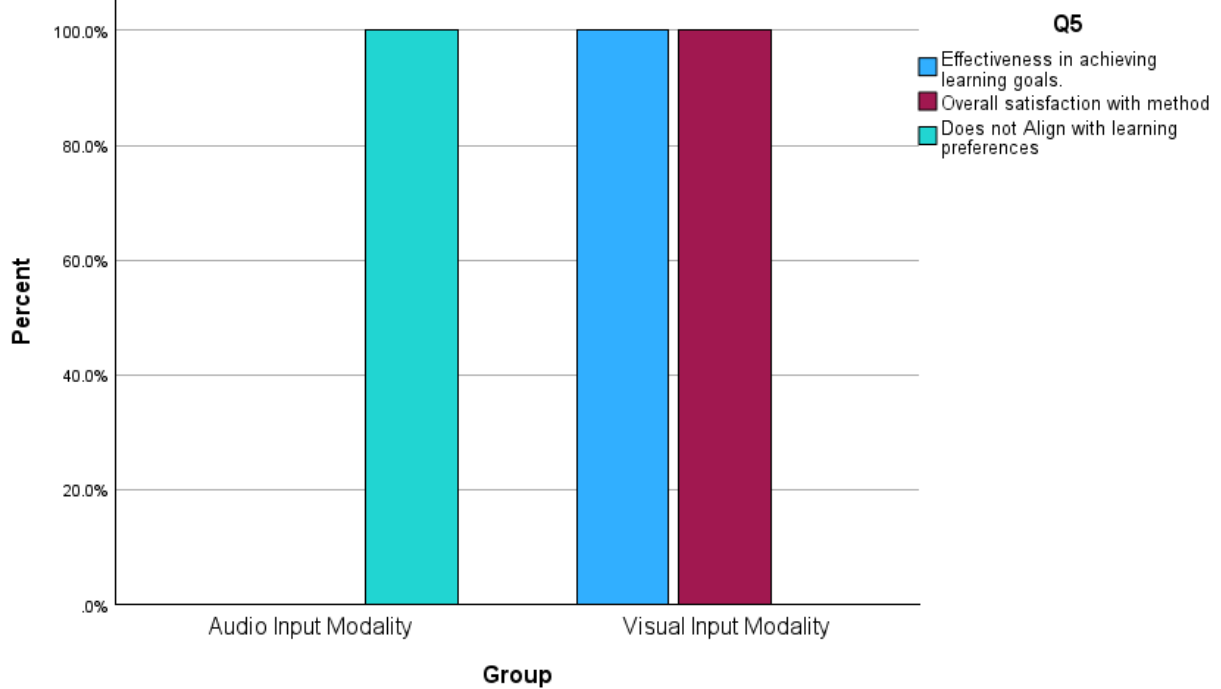
In Figure 5, students reflected on their overall experience with the learning methods. V-group expressed overall satisfaction with their method, with 100% agreement in terms of vocabulary acquisition. By contrast, A-group collectively agreed that this method did not align with their learning preferences, preferring either the visual method or a mixed visual-auditory approach.

Figure 6

Recommendation of Instructional Method

6. Would you recommend the use of the audio/visual-based instructional method based on your experience in this study? Please explain your recommendation.
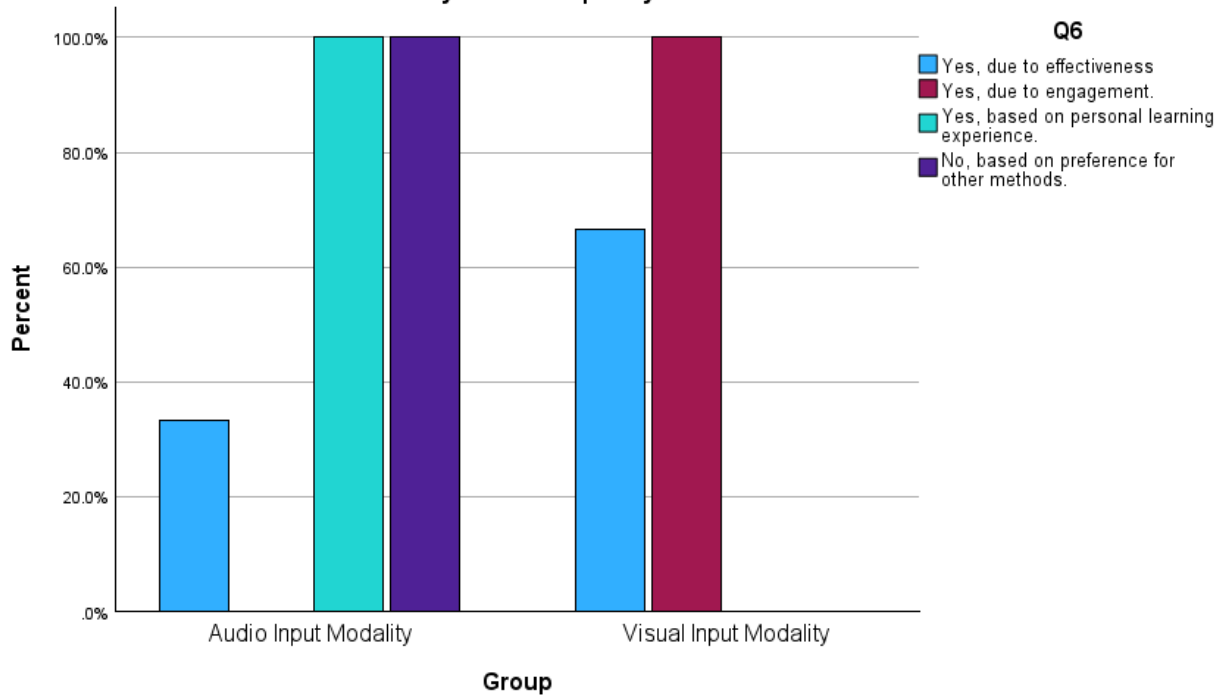
Figure 6 presents the responses to whether students would recommend their assigned learning method. V- group had two main answers: Yes, due to effectiveness (65%), and yes, due to engagement (100%). The A-group had three primary ones: No, preferences for other methods (100%), yes, personal learning experience (100%), and yes, effectiveness (35%). Though indicated a preference for the auditory method based on personal experience, 3 out of 4 students from the auditory group preferred and recommended the visual method.

2. Likert Scale Questions

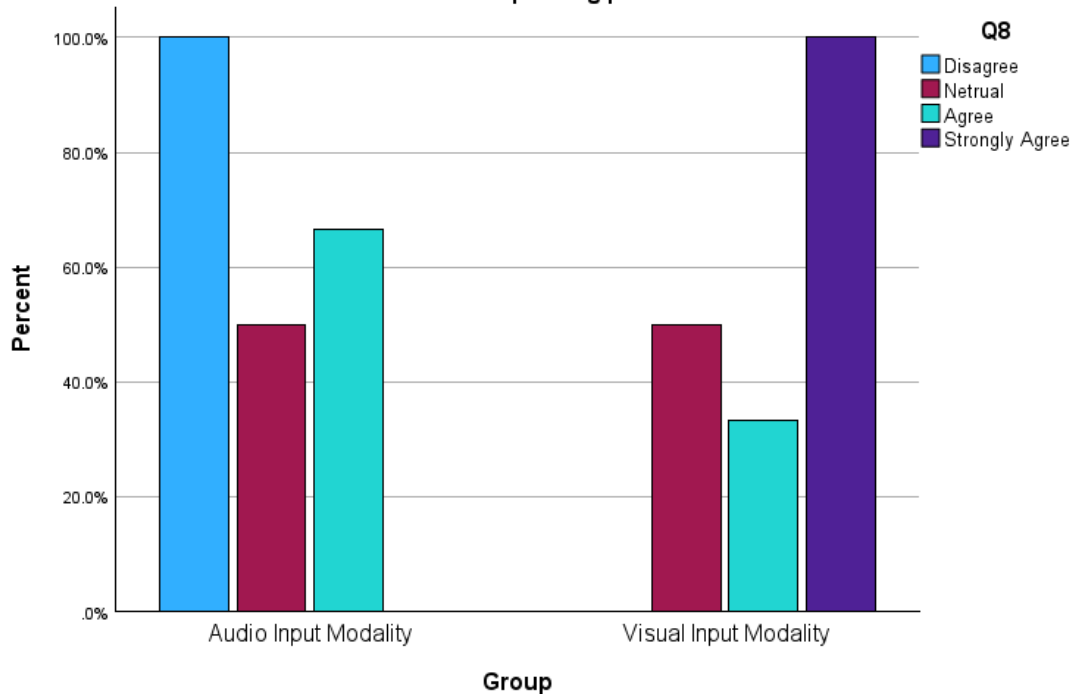**Figure 7**

Vocabulary Acquisition

7. Rate your level of agreement with the statement: "The audio/visual-based instructional method helped me improve my vocabulary acquisition."
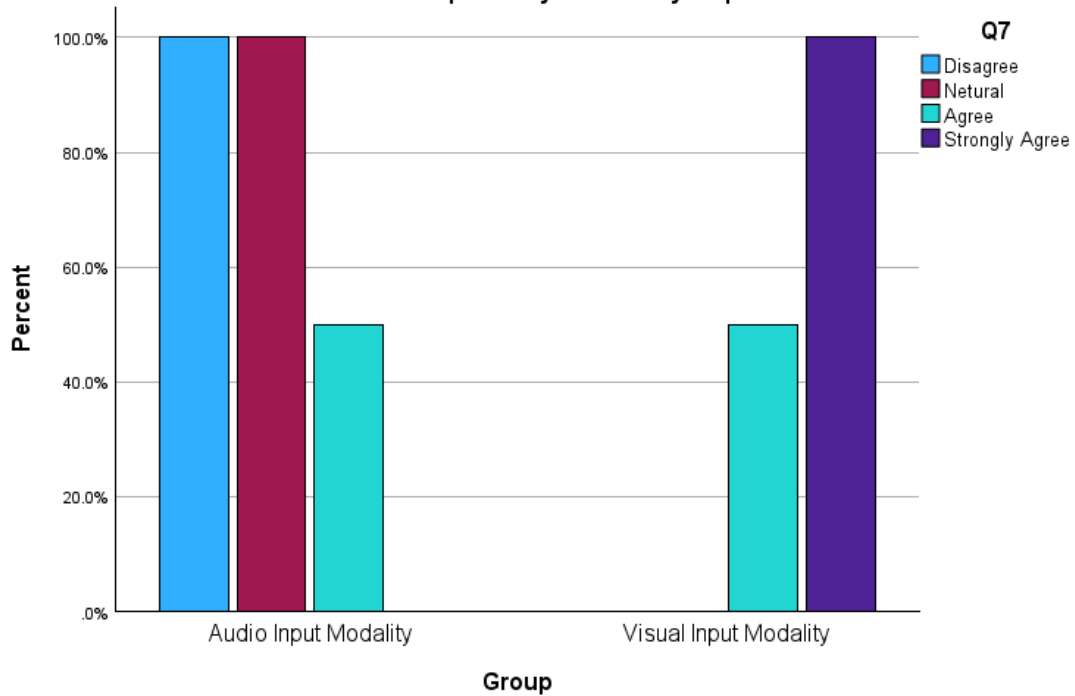
Moreover, in Figure 7, students rated their level of agreement with the statement "The Audio/visual-based instructional method helped me improve my vocabulary acquisition." While A-group's responses were (Disagree %25, Neutral %25, & Agree %50), V-group were (Agree %50 & Strongly Agree %50).

Figure 8

Speaking Practice

8. Rate your level of agreement with the statement: "The audio-based instructional method facilitated better speaking practice."

In Figure 8, students rated their level of agreement with the statement: "The audio/visual-based instructional method facilitated better speaking practice." A-group's responses were (Disagree %25, Neutral %25, & Agree %50), and V- group were (Neutral %25, Agree %25 & Strongly Agree %50). This can be attributed to the visual aids' impact in enhancing retention and memory through remembering sentence structure, organizing thoughts, and paying attention to contextual cues.

5. Discussion

This study showcases the difference in ESL vocabulary acquisition between visual and auditory learning approaches. Despite suggested similar outcomes from previous studies, the empirical evidence in our study shows the advantage that visual methods hold over auditory ones. Statistical tests and qualitative feedback both demonstrate the potency of visual aids in facilitating comprehension and vocabulary learning. Moreover, not only did the feedback express preference for visual aids, but there was a clear superiority in the structure of written responses from students in the visual group. This aligns with previous literature on the importance of multimodal learning strategies, where

visual cues complement auditory input to get the best results (Valentini et al., 2018). For example, students in V-group often exhibited clearer expressions and more comprehensive thoughts due to richer vocabulary. This can be seen through their use of statements like *" My experience with the visual-based is very effective. It enhances my comprehension to what I am working, and improve pronunciation and accent.".* In contrast, students from A-group conveyed their thoughts of the method using statements such as *" No, I like to be a video because I prefer to not the words. Not are the words hear it clearly. No, I don't prefer to hear audio because it is not clear."* This shows that A-group's responses tended to be less coherent and repetitive. Thus, how differently students use vocabulary, grammar, and structure their sentences is evident between the two groups. While students from the visual group wrote more coherent, linguistically sophisticated, and rich responses, audio group students struggled with clearly articulating their thoughts. Even though the amount of time given to this research study was relatively modest, it demonstrates the deeper understanding of vocabulary through visual stimuli, and limitation through auditory means alone.

Despite errors in the answers of both groups, visual group students showed a higher level of language proficiency overall. This indicates that errors are unavoidable during the learning process, but the visual method facilitates a more comprehensive vocabulary understanding, subsequently leading to improved speaking, and writing skills over time (Mohd Tahir et al., 2020). The lack of statistically significant differences in pre-test results between the two groups indicates that initial proficiency levels did not influence the effectiveness of the visual learning method. Instead, the observed improvements in speaking and vocabulary skills among visual learners throughout the duration of the ESL course highlight the adaptive nature of visual learning strategies and their capacity to accommodate learners of varying proficiency levels.

6. Conclusion

According to Van Der Merwe (2019), effective pedagogy in ESL encompasses and links a number of techniques and approaches, including reflective teaching practices, cognitive processing strategies, specialized training for ESL teachers, adaptive teaching methods, technological incorporation, and incidental teaching. The aforementioned

processes can be linked together, aiding in developing the most effective ways to facilitate vocabulary learning and memorization. Educators can partake in reflective practices to improve teaching methods, leveraging cognitive processes to understand how their students process the information and perceive it.

Furthermore, current studies highlight the importance of incorporating technology into ESL instructional methods, shedding light on the role of different cognitive processes in addition to adaptive learning methods. However, there remains a gap in understanding the impact of visual and auditory input modalities on developing a plethora of vocabulary among students. This paper addresses this gap by investigating the effectiveness of visual and auditory modalities in enhancing vocabulary acquisition through experimental research.

This study reveals the superiority of visual learning methods over auditory ones in ESL vocabulary acquisition and speaking skills development. The efficacy of visual approaches can be seen through our quantitative results and qualitative feedback, where students favored using visual aids due to their enhanced comprehension and engagement. Additionally, all students from the audio group conveyed interest in switching to visual learning due to benefits such as observing hand gestures, facial expressions, and connecting voices to speakers. The group also highlighted the importance of active engagement and not feeling bored during the process, a factor missing in audio-based learning. Despite the small sample size used in this research study, our findings underline the potential of visual-based learning in improving vocabulary acquisition outcomes in ESL settings, assisting educators to create engaging learning environments, tailored to each student or groups' needs. Future research could focus on additional factors that can affect the learning process, such as gender, age, learning disorders, and mental health issues.

Authors' bio

Dr. Rabih Nabhan is an associate professor of Linguistics and Applied Linguistics at Holy Spirit University of Kaslik (USEK). He teaches graduate and undergraduate courses of linguistics to Language, Literature, and Translation students. Currently, Dr. Nabhan is the Director of USEK Language Center and Director of USEK Division of Academic Foundations. His publications emphasize Language Acquisition, Discourse Analysis Learning Disabilities, and Writing Skills.

Joe Khoueiry is an English instructor at the University of Balamand and at the Arab Open University. He has a profound interest in applied linguistics and has dedicated his research to enhancing the linguistic abilities of ESL students. His innovative work includes developing AI-based interventions to aid children with ADHD and dyslexia. Joe has actively participated in multiple seminars on these subjects, demonstrating his commitment to the field. His ongoing research and forthcoming publications underscore his commitment to advance our understanding of linguistics.

References

Afzal, N. (2019). A Study on Vocabulary-Learning Problems Encountered by BA English Majors at the University Level of Education. SSRN Electronic Journal. https://doi.org/10.2139/ssrn.3465990

Akpan, E. E., Fcicn, A., Ppgdca, P., & Lion, C. J (2023, December). Independent T-Test statistics: its relevance in educational reserach.

Al-Homoud, F. A., & Osman, S. N. (2015). Flash Cards vs. Word Lists for Vocabulary Building among Majmmah High School Students. Journal of Human and Administrative Sciences, 7, 1–19. https://doi.org/10.12816/0024010

Alghamdi, H. H. (2018, December 11). Exploring Second Language Vocabulary Learning in ESL Classes. *English Language Teaching*, *12*(1), 78. https://doi.org/10.5539/elt.v12n1p78

Bariuad, S. (2022, June 23). Spaced Repetition: How to Make Your Training Unforgettable. EdApp Microlearning. https://www.edapp.com/blog/how-spaced-repetition-works/

Bin Mohd Saad, Zulfadli & Mangaleswaran, & Yusof, Syahidah & Baharudin, Siti & Hashim, Aneeis & Osman, Tengku & Noor, & Hanim, Alia & Hassan, Mohamad. (2022). 计算机集成制造系统 ISSN No.11 Computer Integrated Manufacturing Systems. 10.24297/j.cims.2022.12.06.

Chaudhari, A. (2021, October 1). What is a True Experimental Design? - Voxco. Voxco. https://www.voxco.com/blog/true-experimental design/#:~:text=The%20true%20experiment%20uses%20random,the%20influence%20of%20independent%20variables.

El Khairat, M. (2021, April). Advancing Vocabulary Acquisition by Integrating Visual Image and Texts in Writing. In International Conference on Applied Science and Technology on Social Science (ICAST-SS 2020) (pp. 493-498). Atlantis Press.

Feng, Y., & Webb, S. (2019, October 18). LEARNING VOCABULARY THROUGH READING, LISTENING, AND VIEWING. Studies in Second Language Acquisition, 42(3), 499–523. https://doi.org/10.1017/s0272263119000494

Ganapathy, M., & Seetharam, S. A. (2016). The Effects of Using Multimodal Approaches in Meaning-Making of 21st Century Literacy Texts Among ESL Students in a Private School in Malaysia. Advances in Language and Literary Studies, 7(2), 143-155.http://dx.doi.org/10.7575/aiac.alls.v.7n.2p.143

Harrison, S. (2023, August 2). English Words - How Many is Enough? - UCT English Language Centre. UCT English Language Centre.

Hien, L. T. N. (2021, June 8). Communicative Language Teaching in Teaching ESL for University Students. *Journal of English Language Teaching and Applied Linguistics*, *3*(6), 49–57. https://doi.org/10.32996/jeltal.2021.3.6.7

Hilgendorf, S. K. History of Language-Teaching Methods. 1-5. https://doi.org/10.1002/9781405198431.wbeal0510.pub

Hill, A. C. (2022, March 16). The Effectiveness of Mnemonic Devices for ESL Vocabulary Retention. English Language Teaching, 15(4), 6. https://doi.org/10.5539/elt.v15n4p6

Jabri, U., & Samad, I. S. (2021). The Implementation of Communicative Language Teaching (CLT) in Teaching English. Majesty Journal, 3(1), 1-7. https://doi.org/10.33487/majesty.v3i1.1186

Kacani, L., & Cyfeku, J. (2015, December 15). Developing EFL Vocabulary through Speaking and Listening Activities. Academic Journal of Interdisciplinary Studies. https://doi.org/10.5901/ajis.2015.v4n3s1p390

Kottacheruvu, N. (2023). English Vocabulary Enhancement Strategies for ESL Learners: A Study. International Journal of Language and Literary Studies, 5(2), 37–54. https://doi.org/10.36892/ijlls.v5i2.1257

Mohd Tahir, M. H., Safinas Mohd Ariff Albakri, I., Mohd Adnan, A. H., Syafiq Ya Shaq, M., & Mohamad Shah, D. S. (2020). The Application of Visual Vocabulary for ESL Students' Vocabulary Learning. SSRN Electronic Journal. https://doi.org/10.2139/ssrn.3649330

Mouritsen, M. L., Davis, J. T., & Jones, S. C. (2016). ANOVA Analysis of Student Daily Test Scores in Multi-Day Test Periods. Journal of Learning in Higher Education, 12(2), 73-82.

Msw, E. H. (2022, May 27). *Understanding Elaborative Rehearsal in Psychology*. Verywell Health. https://www.verywellhealth.com/elaborative-rehearsal-a-better-way-to-memorize-98694

Nakata, T., & Elgort, I. (2020, June 3). Effects of Spacing on Contextual Vocabulary Learning: Spacing Facilitates the Acquisition of Explicit, But Not Tacit, Vocabulary Knowledge.

Second Language Research, 37(2), 233–260. https://doi.org/10.1177/0267658320927764

Olson, J., McAllister, C., Grinnell, L., Gehrke Walters, K., & Appunn, F. (2016, January 4). Applying Constant Comparative Method with Multiple Investigators and Inter-Coder Reliability. The Qualitative Report. https://doi.org/10.46743/2160-3715/2016.2447

Rahmat, N. H., & Mohandas, E. S. (2020, November 19). VOCABULARY ACQUISITION AMONG ESL LEARNERS: A LOOK AT STRATEGIES AND BARRIERS. European Journal of English Language Teaching, 6(2). https://doi.org/10.46827/ejel.v6i2.3431

Technology in ESL Classrooms: A Modern Approach - Bay Atlantic University - Washington, D.C. (2024, March 15). Bay Atlantic University - Washington, D.C. https://bau.edu/blog/technology-in-esl-classrooms/

Tian, F., Li, H., Tian, S., Yang, J., Shao, J., & Tian, C. (2020, June). Psychological Symptoms of Ordinary Chinese Citizens Based on SCL-90 During the Level I Emergency Response to COVID-19. Psychiatry Research, 288, 112992. https://doi.org/10.1016/j.psychres.2020.112992

Valentini, A., Ricketts, J., Pye, R. E., & Houston-Price, C. (2018, March). Listening While Reading Promotes Word Learning from Stories. Journal of Experimental Child Psychology, 167, 10–31. https://doi.org/10.1016/j.jecp.2017.09.022

Van der Merwe, E. (2019). Strategies for Enhancing Vocabulary Acquisition in the Language of Instruction (Doctoral dissertation, North-West University (South Africa)).

Zhang, P., & Graham, S. (2020, June 4). Learning Vocabulary Through Listening: The Role of Vocabulary Knowledge and Listening Proficiency. Language Learning, 70(4), 1017–1053. https://doi.org/10.1111/lang.12411.

Appendix

V0	ID	Student's ID
V1	Group	V1a. Audio V1b. Visual

V2	Question 1: Can you describe your experience with the audio/visual-based instructional method used in this study?	V2a. Engagement V2b. Effectiveness V2c. Clarity of instructions V2d. Enjoyment V2e. Concentration
V3	Question 2: What aspects of audio/visual-based instruction do you find most effective for learning vocabulary and improving speaking skills?	V3a. Pronunciation V3b. Vocabulary acquisition V3c. Contextual examples V3d. Clear explanations V3e. Real-life applications
V4	Question 3: Have you encountered any challenges or difficulties with the audio/visual-based instructional method? If so, please explain.	V4a. Technical issues V4b. Understanding accents V4c. Lack of interactivity V4d. Lack of Concentration V4e. Time constraints and fast speaking
V5	Question 4: How did the use of auditory/visual input impact your learning process?	V5a. Improved listening skills V5b. Increased retention V5c. Enhanced comprehension V5d. Confidence building V5e. Adaptation to auditory learning style
V6	Question 5: Reflecting on your overall experience, do you think the audio/visual-based instructional method was suitable for vocabulary acquisition and speaking improvement? Why?	V6a. Enhanced engagement V6b. Real-world relevance V6c. Effectiveness in achieving learning goals. V6d. Overall satisfaction V6e. Does not Align with learning preferences.
V7	Question 6: Would you recommend the use of the audio/visual-based instructional method based on your experience in this study? Please explain your recommendation.	V7a. Yes, due to effectiveness. V7b. Yes, due to engagement. V7c. Yes, based on personal learning experience. V7d. No, due to limitations. V7e. No, based on p4reference for other methods.
V8	Question 7: "Rate your level of agreement with the statement: 'The audio/visual-based instructional method helped me improve my vocabulary acquisition.'	Strongly Agree: "5" Agree: "4" Neutral: "3" Disagree: "2" Strongly Disagree: "1"
V9	Question 8: Rate your level of agreement with the statement: "The audio-based instructional method facilitated better speaking practice."	Strongly Agree: "5" Agree: "4" Neutral: "3" Disagree: "2" Strongly Disagree: "1"