Journal of Al-Farabi for Humanity Sciences Volume (8), Issue (1) July(2025)

ISSN: 2957-3874 (Print)
Journal of Al-Farabi for Humanity Sciences (JFHS)
https://iasj.rdd.edu.iq/journals/journal/view/95
مجلة الفارابي للعلوم الإنسانية تصدرها جامعة الفارابي

Investigating the Influence of Multimodal Approaches on EFL University Students' Audio-Visual Comprehension

Ibrahim Hassan Ali[1] (M.A and Ph.D. Student at Tikrit University) - Corresponding Author
Anbar University, University Headquarter,
Division of Psychological Counseling and Educational Guidance, Ramadi, IRAQ.
ORCID ID: 0009-0004-9369-6488
ibrahim.hasan@uoanbar.edu.iq
Istabraq Tariq Jawaad Alazzawi[2] (Ph.D.)
University of Tikrit, College of Education for Humanities,
Dept. of English, Tikrit, IRAQ.
ORCID ID: 0000-0002-5590-343X
astbraktarek@tu.edu.iq

دراسة تأثير المناهج متعددة الوسائط على فهم طلاب الجامعات الناطقين باللغة الإنجليزية كلغة أجنبية للمواد السمعية والبصرية

إبراهيم حسن علي / جامعة الأنبار
إستبرق طارق جواد العزاوي / جامعة تكريت، كلية التربية للعلوم الإنسانية، قسم اللغة الإنجليزية

Abstract

This study explores how different multimodality affects audio-visual comprehension in English as a Foreign Language (EFL) university students. The study included a sample of 214 students from Anbar and Tikrit Universities. To measure the effectiveness of different methods, a diagnostic test is used as the main tool to gauge the comprehension levels of students who were exposed to multimodal learning methods (visual, auditory and textual stimuli). The results indicate some level of audio-visual comprehension among students that read light and heavy multimodal material. It underscores the role of the multimodal approach in developing EFL learners' comprehension skills and its potential value for wider implementation in language teaching.**Keywords:** Multimodal approach, EFL, audio-visual comprehension, language learning, university students.

المستخلص

تستكشف هذه الدراسة كيف تؤثر الوسائط المتعددة المختلفة على الفهم السمعي البصري لدى طلاب الجامعات الذين يتعلمون اللغة الإنجليزية كلغة أجنبية، شملت الدراسة عينة مكونة من ٢١٤ طالبًا من جامعتي الأنبار وتكريت. تم استخدام اختبار تشخيصي كأداة رئيسية لتقييم مستويات الفهم لدى الطلاب الذين تعرضوا لأساليب التعلم متعددة الوسائط (المحفزات البصرية والسمعية والنصية). تشير النتائج إلى وجود مستوى معين من الفهم السمعي البصري بين الطلاب الذين قرأوا موادًا متعددة الوسائط خفيفة وثقيلة. كما تؤكد النتائج على دور المنهج متعدد الوسائط في تطوير مهارات الفهم لدى متعلمي اللغة الإنجليزية كلغة أجنبية وقيمته المحتملة للتطبيق الأوسع في تعليم اللغة الانكليزية.

الكلمات المفتاحية: المنهج متعدد الوسائط، اللغة الإنجليزية كلغة أجنبية، الفهم السمعي البصري، تعلم اللغة، طلاب الجامعات.

Introduction

The emphasis on multimodal approaches in the field of English as a Foreign Language (EFL) education within contemporary pedagogy presents opportunities to improve language learning outcomes. A multimodal approach employs different methods of communication — visual, auditory, textual, and kinesthetic elements—to promote

understanding and retention (Kress & van Leeuwen, 2001). Since, the access of digital media is growing, educators are adopting many audio-visual materials (videos, podcasts, interactive multimedia, etc.) to EFL classrooms to develop student's listening and viewing comprehension. Additionally, it discusses the development of English as a Foreign Language (EFL) university students' audio-visual comprehension through a multimodal approach and the importance of implementing diverse sensory inputs simultaneously in the acquisition of a language as well as the implications of engagement and cognitive processing in the process.Multimodality refers to the use of multiple semiotic resources such as images, sounds, and gestures for meaning making (Jewitt, 2006), and derived its theoretical base from social semiotics. In the language education domain, multimodal pedagogy is in line with Mayer's Cognitive Theory of Multimedia Learning (2005), which argues that opportunities for information presentation through visual and auditory modalities are more beneficial for learners than modalities presented separately. Dual coding—processing information both verbally and visually—is said by Mayer to alleviate cognitive load, and contribute to retention.Multimodal input in EFL learning exposes learners to authentic language use, as it reflects communication in the real world. Plass and Jones (2005) have conducted research that has shown that students showed better comprehension and vocabulary learning when they learned through multimedia-based instruction instead of text-only. Multimodal materials also address different learning styles presented by students, where visual, auditory and kinesthetic learners are involved (Fleming & Mills, 1992).Audio-visual comprehension is an essential part of language proficiency, which refers to the ability to comprehend literacy by listening to spoken language in the presence of accompanying visual clues (e.g., gestures, facial expressions, contextual landscapes) (Wagner, 2010). Because of limited study abroad and exposure to natural speech, accents, and conversational speed, however, EFL learners struggle with developing this skill (Vandergrift & Goh, 2012). And traditional listening exercises provide enough contextual support to frustrate students or keep them from wanting to pay attention.But multimodal resources like films, news clips and educational videos provide a fuller linguistic and paralinguistic environment to support comprehension. For instance, Canning-Wilson (2000) and Baltova (1999) provide evidence that visual support of videos facilitates learners in inferring meaning, recognizing patterns of intonation and understanding the cultural context. In addition, the interaction of subtitles and images has been shown to assist in the recognition and retention of words (Guichon & McLornan, 2008).A few empirical studies provide evidence on the effectiveness of multimodal methods to improve EFL learners audio-visual understanding. Suvorov (2015) found that students who viewed video content with captions, scored higher in listening comprehension than students who viewed without captions. Intermedia perspectives on L2 input can help uncover some of these possibilities; as Sydorenko (2010) found, learners learned more vocabulary and idiomatic expressions by using content such as videos with subtitles or videos with interactive transcripts.The use of multimedia platforms such as YouTube and Ted Talks, for example, has been shown to increase motivation and student engagement (Watkins & Wilkins, 2011). Examples of these interactive multimedia components imported from other disciplines include language learning apps (i.e. Duolingo, Rosetta Stone) which incorporate gamification and visual storytelling; enhancing retention by repeated visibility and contextual reinforcement of concepts (Godwin-Jones, 2018).However, there is a drawback to adopting multimodal features in EFL classrooms. An important consideration touching all areas of a digital learning experience is cognitive overload (Sweller, 2011). To counteract this, educators should strategically choose and scaffold multimedia materials to better match learners' ability levels (Huang & Eskey, 2000).Another challenge is the digital divide—the unequal access to technology—which may limit some students' exposure to multimodal learning resources (Warschauer, 2003). To ensure equitable learning opportunities, institutions must address these infrastructural barriers. The research objectives and significance are as follows:

1. Investigate how multimodal approaches influence EFL university students' audio-visual comprehension.
2. Examine the differences among different multimodal ensembles (modes) of EFL university students' performance in audio-visual comprehension test

These discoveries will add up to the corpus of studies on tech-enhanced language acquisition, providing additional for curriculum creators and educators attempting to maximize multimedia usage in EFL education. Incorporating audio-visual resources for a new paradigm for EFL education. This approach is backed by cognitive learning theories and is effective for the diverse range of language learner needs found in the classroom as it caters to both visual and auditory stimulus. Despite the socio-technological challenges of cognitive overload and access that remain, empirical evidence suggests that multimodal resources have enhanced

audio-visual comprehension. Future studies may need to address retention over time and the possible effects of a new trend of technologies such as virtual reality on EFL learning outcomes."

2. Literature Review

2.1 Multimodal Digital Language Learning

"Vodcasts" are recordings of videos that are disseminated over web feeds. Factors such as the proliferation of computers with multimedia capabilities, improvements in broadband technology that allow for large media file downloads, the simplification of the downloading and streaming processes, and the relative ease of discovering new technologies have contributed to their meteoric rise in popularity in recent years (Campbell, 2005). A number of factors have contributed to the meteoric rise of podcasts, including the proliferation of editing software like Odeo and Audacity and distribution sites like iTunes, Education Podcasts, and podOmatic. Vodcasts are a type of mobile learning (m-learning), according to Evans (2008), because of its many uses. "Learn across various contexts through social and content interactions via personal electronic devices" (Crompton, 2013, p.4) is what they enable students to do.As technology improves and the need to make language learning more accessible grows, multimodal digital learning tools like podcasts are finding more and more uses in traditional language courses. The term for this cutting-edge method is blended language learning (abbreviated BLL). Stracke states that BLL is "a specific learning and teaching environment that integrates face-to-face (f2f) instruction with computer-assisted language learning (CALL)" (2007, p.57). Additionally, Figure 1 shows how blended language learning helps students learn by combining traditional in-person teacher-student interaction with technology-enhanced language assignments.

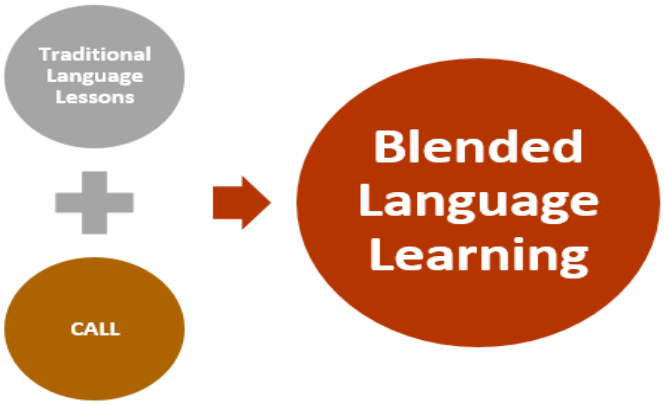


Figure 1.

Blended language learning has several components.

Face to face language training is traditional, no tech used and a teacher in front of students. Such interaction as a form of language pedagogy has been organized for hundreds of years, and many studies have also investigated significant features of this interaction in terms of the social negotiation of meaning, the overall effect of both interactive and non-interactive courses, and the consequences of such instructional methods on students' language learning (Morell 2004). The blending of technology-based learning with face to face lessons allows the process of acquiring knowledge to be a richer experience, where the teaching and learning opportunities arising from multimodal communication are essential. Many researchers (Chinnery, 2006; Kukulska-Hulme, 2013) have introduced a new way of learning languages by computers, referred as Mobile-Assisted Language Learning (MALL). Mobile Augmented Language Learning (MALL) is defined as "mobile technologies in language learning, in particular when the portability of devices indeed offers particular advantages" (Kukulska-Hulme, 2013, p.3701). MALL has more in common with CALL, mobile learning, which emphasizes learning anytime and everywhere with electronic devices, and second language acquisition than with the two. The use of "personal, portable, flexible" devices that made it easy to learn and were accessible "without barrier" and "at all times" (Kukulska-Hulme and Shield 2008) makes the fundamental distinction between Mobile-Assisted Language Learning (MALL) and Computer Assisted Language Learning (CALL).

2.2 EFL AUDIO-VISUAL COMPREHENSION

There are five modalities of information input currently used in hearing assessments: audio-only input and four visual modalities, which are context-only stills, context-only video, content stills and content video (Ockey, 2007). As Ockey postulates, if the input procedures of the hearing exam change, it would create differences in the way in which test-takers process input, likely leading to differences in performance on the test and subsequently impacting the construct validity of the listening assessment. As a result, most related studies focus

on measuring the effects of different audiovisual input appliances on test results and how test takers perform (Ockey, 2007). Previous research has largely compared the effects of audio-only input with video captures. Rajabi et al. (2021) conducted a study with 91 second language learners that found no significant differences in student performance between audio-only and video-mediated assessments. He also discovered that some students, seemingly distracted by visual distractions, made the decision not to see the television. Open-ended examination questions were used to explore the audio–visual divide among 104 learners of English in Hong Kong (Coniam, 2001). The results suggested the audio-only group outperformed the video group; however, this difference was not statistically significant. Furthermore, the video group and the audio group did not consider the utilization of video as a medium helpful in understanding when listening. In contrast, 36% of examinees reported not looking at the screen at all during the assessment, and only a small percentage reported finding the video distracting. Cubilo and Winke (2013), using writing and note-taking tasks to assess listening comprehension found that quality of writing post-listening remained stable under conditions of visual input and auditory input, while note-taking behaviour of participants demonstrated differences, specifically that quality of note-taking decreased under visual condition significantly. In contrast, Wagner (2010) reported that the video group performed 6.5% better than the audio group on the post-test, where this difference was statistically significant. He thinks that the video's nonverbal clues help the subjects improve. For example, Sueyoshi and Hardison (2005) examines the role of speakers' facial expressions and gestures embedded within videotexts on L2 learners' listening comprehension. Forty-two low-intermediate and advanced English learners were randomly assigned to view one of three modalities of the stimuli: audio-visual with gesture and facial expression, audio-visual without gesture, or audio only. Sueyoski and Hardison discovered that visual input improved student performance at both English proficiency levels. Higher academic level students performed significantly better in the audio-visual condition with a face, while lower academic level students achieved better performance with gesture + face input. According to the questionnaire data, students were positive about visual cues, especially during in-person encounters.

3. Methodology

3.1 Research Design

Descriptive design in educational research is a methodological approach that aims to systematically characterize, document, and analyze educational phenomena, populations, or situations without manipulating variables. This research methodology focuses on answering questions about who, what, where, when, and to what extent, rather than explaining why something occurs. Descriptive research in education serves as a fundamental component of the scientific process by helping researchers observe educational phenomena and generate hypotheses based on these observations (Siedlecki, 2020; Blessing et al., 1998). A diagnostic test in the educational context is a type of assessment administered at the beginning of a learning experience—such as a new unit, lesson, or course—to determine students' existing knowledge, skills, strengths, and learning gaps related to a specific subject or topic. Unlike graded or summative assessments, diagnostic tests are typically low-stakes and not used for assigning grades. Instead, their primary purpose is to provide teachers with detailed insights into what students already know and what misconceptions or difficulties they may have, enabling educators to tailor instruction, address specific needs, and plan effective remedial or enrichment activities (Bejar, 1984; Tweed & Wilkinson, 2012; Leighton & Gierl, 2007).

3.2 Participant

The study population of quantitative method consist of third year EFL university students enrolled in the English department of college of education from three universities (i.e. Anbar university and Tikrit university for both college of education for humanities and college of education for woman) during the academic year 2023-2024. The total population 214 students. The quantitative data method of the study was established by convenience sampling. The justification of the using the convenience sampling method is to identify the nearest and easily accessible sample that the researcher can get. Those 214 students represent the sample who are asked to answer verbal and non-verbal communication questionnaire scale. Those students had the same educational and language background knowledge.Table 1. The Dimorphic data on the study participant of the quantitative dimension

Variable	Introductory Features	N	%
population and Sample	Collage of education for humanities university of Tikrit	87	40.65%
	College of education for woman, university of Tikrit	54	25.23%
	College of education for humanities Anbar university	73	34.11%
Gender	Female	153	71.49%
	Male	61	28.50%
Age	19-21	178	83.17%
	22-24	36	16.21%

3.2 Data collection instrument

In this study, a diagnostic test was employed as the primary data collection instrument to assess the influence of multimodal approaches on EFL university students' audio-visual comprehension. The diagnostic test is designed to measure students' weakness and strength comprehension levels when exposed to multimodal learning materials (e.g., videos, images, subtitles, and audio cues). The test items included multiple-choice questions, short-answer responses, and true/false statements based on audio-visual content to evaluate different levels of comprehension. Validity indicates how well an instrument measures something the way we intend it to (Cohen et al., 2018). Also several validation strategies were applied to make sure the diagnostic test was valid. To begin with, a group of EFL experts reviewed the instrument for content validity and comprehensibility to establish face validity. A pilot study was also conducted with some students to finalize the test before implementation. Second, the test items were examined by two EFL specialists in multimodal learning and language assessment to confirm their content validity, indicating that they were representative of audio-visual comprehension and fitting with the intention of the current study (Taherdoost, 2016; Creswell & Creswell, 2018). This quantitative data obtained through the diagnostic test enabled a more refined examination of EFL learners' development in the area of audio-visual comprehension, which also serves to strengthen the multimodal validity of the analysis (Bryman, 2016).The reliability–that is, whether the test gives the same result when the test is repeated – (Field, 2018) were also carefully analyzed. The internal consistency was described by Cronbach's alpha, which quantified the internal coherence of multiple-choice and short-answer items. An acceptable reliability (Tavakol & Dennick, 2011) was achieved with a coefficient of 0.80. In addition, test-retest reliability was assessed by administering the same test to a sub-sample of participants (n = 15) two times, with an interval of two weeks between administrations. Pearson's correlation is calculated to determine the stability of scores and a coefficient of 0.732 is obtained, which is within the acceptable range of reliability (Kline, 2015; McHugh, 2012).

3.3 Procedures of Data collection

The vodcasts used in the study are named English is GREAT, and they can be downloaded for free from iTunesU. English is GREAT (part 1) is part of a series of videocasts. Britain is fantastic. The vodcast lasts 5 minutes and 18 seconds and demonstrates how the English language has evolved over time. Richard, the presenter, visits the British Library in London to interview Roger Walshe, the Head of Learning. The vodcast tells viewers about the library's contents, the impact of the Industrial Revolution, the Internet's influence on English, and the language's adaptability. The British Council's vodcasts were chosen for two reasons. The first was the organization's well-known reputation in language instruction. Second, the vodcasts are linked to the research participants' unique syllabuses. Furthermore, they meet some of the criteria stated by Rosell-Aguilar (2007) to consider when selecting vodcasts for language acquisition. Finally, vodcasts used a range of semiotic modalities of communication (such as still images, gestures, written language, spoken language, and music) to aid in the formation of meaning.

The data collected through audiovisual comprehension tests were analyzed using quantitative data analysis procedures. Numerical statistical analysis was performed using the Statistical Package for the Social Sciences (SPSS), Version 26. A one-sample t-test was conducted, and a one-way Analysis of Variance (ANOVA) was employed to examine differences in comprehension modes among the study sample.

4. Findings and data analysis

4.1.1 Identifying the Difference Between Theoretical Mean and EFL university students' Audio-visual comprehension The first hypothesis aimed to determine whether there was a significant difference between the theoretical mean and EFL university students' audio-visual comprehension scores. The students' mean score on the audio-visual comprehension test was 67.59, compared to the theoretical mean of 50, with a standard deviation

of 9.90. A one-sample t-test was conducted to analyze the data. The calculated t-value 25.994 exceeded the tabulated t-value 1.972 at a 0.05 significance level with 213 degrees of freedom. As shown in Table 2, the results revealed a statistically significant difference between the theoretical mean and the students' audio-visual comprehension scores, favoring the study sample. These findings suggest that the students demonstrated a strong and statistically significant level of proficiency in the audio-visual comprehension test.

Table 2

The Mean Scores, Standard Deviations, and One Sample T-Value of the Students' Audio-Visual Comprehension Test

Group	No. of students	Mean	SD.	Theoretical Mean	T-Value		DF	Level of Sig.
AVC	214	67.59	9.90	50	Calculated	Tabulated	213	0.05
					25.994	1.972		

4.1.2 The differences among different multimodal ensembles (modes) in AVC test

In order to "finding out the differences among different multimodal ensembles (modes) of EFL university students' performance in AVC test," the one-way analysis of variance is used to obtain differences among three variables, as shown in the following table:

Table 3

One-Way Analysis of Variance (ANOVA) Among Multimodal Ensembles (modes)

Groups	Sum of Squares	DF	Mean Square	F-value		Sig.
				Calculated	Tabulated	
Between Groups	1533.991	2	766.995	95.139	3.01	0.05
Within Groups	5151.519	639	8.062			
Total	6685.509	641				

Table 3, shows that the calculated f-value value 95.139 is higher than the tabulated value 3.01, and DF = 2,639 at the 0.05 level of significance. This indicates that there are significant differences among EFL university students' multimodal ensembles in their performance in AVC test.

Table 4

Differences of Means Among Three Variables (Tukey B[a])

Groups		N	Subset for alpha = 0.05		
			1	2	3
Tukey B[a]	Two modes	214	12.03		
	Three modes	214		13.14	
	Four modes	214			15.72
a. Uses Harmonic Mean Sample Size = 214.					

Tukey formula is used to find out the differences among three variables. The findings show that the mean scores of two modes are 12.03, three modes are 13.14, and four modes are 15.72, with a harmonic mean sample size of 214. This indicates a significant difference in the performance of EFL university students' multimodal ensembles in an AVC test, favouring the four modes.

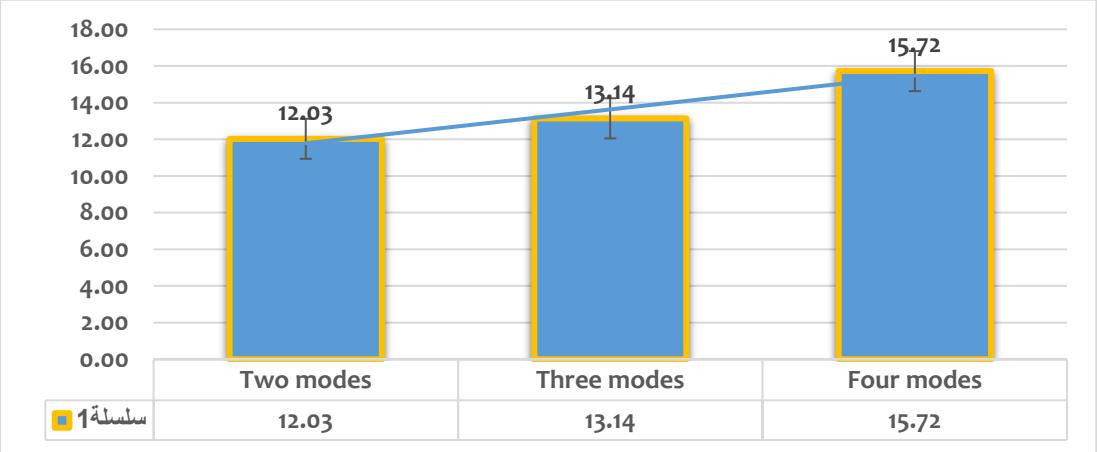
4.2 Discussion of Results

The current study aimed to provide a little additional knowledge about AVC, VC and, NVC and the effect of multimodal visual methodologies supports on EFL environments. The study demonstrated that adopting multimodality had an effective impact on EFL university students at three different colleges of education. The results of the first aim on multimodal visual methodologies and their influence on AVC among EFL university students provide significant insights into how these methodologies enhance learning outcomes over time. The results of the one-sample t-test comparing EFL university students' audio-visual comprehension (AVC) scores to the theoretical mean (50) reveal that students' mean AVC score (67.59) was significantly higher than the theoretical mean (50), with a large t-value (25.994) exceeding the critical t-value (1.972) at $p < 0.05$. This indicates that the difference is statistically meaningful and not due to chance. The higher mean score suggests that the EFL students performed well above the baseline (theoretical) expectation in audio-visual comprehension. The results robustly support the hypothesis that EFL students outperformed the theoretical

mean in AVC. Results according to the second question, indicates a significant difference in the performance of EFL university students in AVC test, based on the number of multimodal ensembles (modes) they were exposed to. Results from the one-way ANOVA and follow-up Tukey post-hoc tests (d) provide insight into how different multimodal presentation configurations impact comprehension outcomes. Overall, the current result indicates statistically significant differences between groups with a different number of modes offered in the comprehension test. The different modes effectiveness may be due to their ability to engage learners through different sensory experiences, therefore facilitating greater understanding and retention. A post-hoc comparison Tukey analysis unpacks these differences, showing how the mean scores compare across the three groups (two modes 12.03; three modes 13.14, four modes 15.72 (Figure 1). The results of the multimodal ensembles indicate that EFL university students' performance on the AVC test can be significantly affected by the number of modes such as depicted in Figure 2 where students exposed to their own four modes performed better than students exposed to fewer options.These findings may be akin to those found in research by Sueyoski and Hardison (2005) and Wagner (2010), which revealed a significant preference for video content. Consider that vodcasts are multimodal audio-visual assets that use a variety of modes to aid in the process of meaning-making. Thus, combining verbal and nonverbal modalities may be effective for improving foreign language understanding. Anxiety-related questions were included to determine whether students felt anxious while trying to understand the messages reflected in the vodcasts. This allowed them to get prior knowledge about the vodcasts. The orchestration of modes in audio-visual content may reduce the difficulty of understanding the message, as well as the amount of tension. The level of fear decreases as the message becomes more understandable. For this reason, teachers must pay more attention to the selection of appropriate materials and teaching methods so that pupils feel more confident and less apprehensive.

Figure 2

Comparison Among Audio-visual Comprehension Modes



5. Conclusion

This study found substantial positive impact of multimodal applications on AVC of EFL university students. The findings show that students who had access to multimodal methodologies always scored better than the hypothetical mean, as highlighted by the high difference in AVC obtained (67.59 vs. 50). This indicates the effectiveness of exploiting different sensorial and representational modes to improve EFL reading environments in terms of comprehension, engagement, and retention. In addition, the study demonstrates how the quantity of multimodal ensembles affects AVC, with students utilizing four distinct modes receiving significantly greater scores than those with three and two modes. This implies that more complex multimodal input (e.g., visual, auditory, textual, etc.) results in deeper understanding, hence more successful learning programs. These results support the cognitive load theory and dual coding theory, coming from the premise that multimodal learning eases information processing and eases cognitive effort spread across various modalities. Overall, based on this research, it can be a valuable empirical evidence in support of multimodal in EFL. They should be aware of the optimal use of the multi-modal methods of representation to maximize the audio-visual understanding ease and the overall language performance. Further studies might focus on the sustainability of the retention effects of multimodal learning and its effectiveness across various levels of proficiency and cultural backgrounds.

5.1 Recommendations

The following recommendations shall be provided for educators, curriculum designers, and researchers if considering the findings of the present study on the positive influence of utilizing multimodal methodologies on AVC of EFL university learners:

1. Integrate Multimodality in Instruction: As the study discovered the type of AVC is increased with exposure to more modes (text, photo, sound, audio,)teachers should plan the lessons in ways that feature visual-auditory and kinesthetic Reinforcing comprehension through similar modalities — like subtitled videos, infographics, and interactive discussions.
2. Utilize Technology-Enhanced Learning Tools: Leverage digital tools (such as interactive whiteboards, language-learning applications, and multimedia content) to offer various input sources and engage students in multimodal learning activities.
3. Gradually Build Complexity: Begin with more straightforward multimodal resources (two modes), then progressively increase the complexity with three or four modes, enabling learners to adapt and gradually enhance their understanding skills.
4. Create Multimodal EFL Material: Mainstream textbooks and instructional assets should embed sound and visual snippets preferably audio, videos, images, and activities to address individualized preferences.
5. Support Delivering Teacher Training on Multimodal Pedagogy: Because successful multimodal learning is an effective pedagogical practice, professional learning programs should inform the people who teach these skills on how to select, design, and effectively implement multimodal content.
6. Incorporate Blended Learning Models: Integrate traditional with online multimedia approaches to provide students with diverse and flexible learning experiences.

5.2 Suggestions for Future Studies

- Investigate whether the benefits of multimodal learning persist over time by conducting longitudinal studies on AVC retention among EFL students.
- Assess how factors such as language proficiency, learning styles, and cultural background influence the effectiveness of multimodal approaches.
- Experiment with emerging technologies (e.g., virtual reality, augmented reality, AI-driven interactive content) to determine their impact on AVC in EFL contexts.

References

Baltova, I. (1999). Multisensory language teaching in a multidimensional curriculum: The use of authentic bimodal video in core French. Canadian Modern Language Review, 56(1), 31-48.

Bejar, I. I. (1984). Educational diagnostic assessment. Journal of educational measurement, 21(2), 175-189.

Blessing, L. T., Chakrabarti, A. M. A. R. E. S. H., & Wallace, K. M. (1998). An overview of descriptive studies in relation to a general design research methodology. Designers: The key to successful product development, 42-56.

Bryman, A. (2016). Social research methods (5th ed.). Oxford University Press.

Campbell, G. (2005). “There is something in the air: Podcasting in education”. EDUCAUSE Review, 40(6), 32-46. Retrieved from http://eric.ed.gov/?id=EJ745798

Canning-Wilson, C. (2000). Practical aspects of using video in the foreign language classroom. The Internet TESL Journal, 6(11).

Chinnery, M. G. (2006). “Going to the MALL: Mobile assisted language learning”. Language Learning and Technology, 10 (1), 9-16.

Cohen, L., Manion, L., & Morrison, K. (2018). Research methods in education (8th ed.). Routledge.

Coniam, D. (2001). “The use of audio or video comprehension as an assessment instrument in the certification of English language teachers: A case study”. System, 29, 1-14. doi:10.1016/S0346-251X (00)00057-9

Coniam, D. (2001). The use of audio or video comprehension as an assessment instrument in the certification of English language teachers: a case study. System 29, 1–14. doi: 10.1016/S0346-251X(00)00057-9

Creswell, J. W., & Creswell, J. D. (2018). Research design: Qualitative, quantitative, and mixed methods approaches (5th ed.). Sage.

Crompton, H. (2013). “A historical overview of mobile learning: Toward learner-centered education”. In Z. Berge & L. Muilenburg (Eds.) Handbook of mobile learning. New York: Routledge, 3-14.

Cubilo, J., and Winke, P. (2013). Redefining the L2 listening construct within an integrated writing task: considering the impacts of visual-cue interpretation and note-taking. Lang. Assess. Q. 10, 371–397. doi: 10.1080/15434303.2013.824972

Evans, C. (2008). "The effectiveness of m-learning in the form of podcast revision lectures in higher education". Computers & Education, 50, 491-498. doi: 10.1016/j.compedu.2007.09.01
Field, A. (2018). Discovering statistics using IBM SPSS Statistics (5th ed.). Sage.
Fleming, N. D., & Mills, C. (1992). Not another inventory, rather a catalyst for reflection. To Improve the Academy, 11(1), 137-155.
Godwin-Jones, R. (2018). Second language vocabulary learning through multimedia. Language Learning & Technology, 22(3), 1-24.
Guichon, N., & McLornan, S. (2008). The effects of multimodality on L2 learners: Implications for CALL resource design. System, 36(1), 85-93.
Huang, H., & Eskey, D. E. (2000). The effects of closed-captioned television on the listening comprehension of intermediate English as a second language (ESL) students. Journal of Educational Technology Systems, 28(1), 75-96.
Jewitt, C. (2006). Technology, literacy, and learning: A multimodal approach. Routledge.
Kline, P. (2015). A handbook of test construction. Routledge.
Kress, G., & van Leeuwen, T. (2001). Multimodal discourse: The modes and media of contemporary
Kukulska-Hulme, A. (2013). "Mobile-assisted language learning". In C. Chapelle (Ed.), The encyclopedia of applied linguistics. New York: Wiley, 3701-3709.
Kukulska-Hulme, A. and Shield, L. (2008). "An overview of mobile assisted language learning: From content delivery to supported collaboration and interaction". ReCALL, 20 (3), 271–289. doi:10.1017/S0958344008000335
Leighton, J., & Gierl, M. (Eds.). (2007). Cognitive diagnostic assessment for education: Theory and applications. Cambridge University Press.
Mayer, R. E. (2005). The Cambridge handbook of multimedia learning. Cambridge University Press.
Mayer, R. E. (2009). Multimedia learning. New York: Cambridge University Press.
McHugh, M. L. (2012). Interrater reliability: The kappa statistic. Biochemia Medica, 22(3), 276–282.
Morell, T. (2004). La interacción en la clase magistral. Alicante: Servicio de Publicaciones de la Universidad de
Ockey, G. J. (2007). Construct implications of including still image or video in computer-based listening tests. Lang. Test. 24, 517–537. doi: 10.1177/0265532207080771
Plass, J. L., & Jones, L. C. (2005). Multimedia learning in second language acquisition. The Cambridge handbook of multimedia learning, 467-488.
Rajabi, P., Mahmoodi, K., and Hosseini, S. A. (2021). Flipped classroom model and its impact on Iranian EFL learners' classroom anxiety and listening performance. Comput. Assist. Lang. Learn. J. 22, 1–16. Available online at: http:// www.callej.org/journal/22-3/Rajabi-Mahmoodi-Hosseini2021.pdf
Ramirez, D. and Alonso, I. (2007). "Using digital stories to improve listening comprehension with Spanish young learners of English". Language Learning & Technology, 11(1), 87-101.
Rosell- Aguilar, F. (2007). "Top of the pods- in search of a podcasting "pedagogy" for language learning". Computer Assisted Language, 20 (5), 471-492. doi:10.1080/09588220701746047
Rost, M. (2002). Teaching and researching listening. Harlow, England: Longman.
Siedlecki, S. L. (2020). Understanding descriptive research designs and methods. Clinical Nurse 34(1), 8-12.
Stockwell, G. and Hubbard, P. 2013. "Some emerging principles for mobile-assisted language learning". Monterey, CA: The International Research Foundation for English Language Education. Retrieved from http://www.tirfonline.org/englishin-the-workforce/mobile-assisted-language-learning.
Stracke, E. (2007). "A road to understanding: A qualitative study into why learners drop out of a blended language learning (BLL) environment". ReCALL 19 (1) 57–78.
Sueyoski, A. and Hardison, D. (2005). "The role of gestures and facial cues in second language listening comprehension". Language Learning, 55, 661-699. doi:doi.org/10.1111/j.0023-8333.2005. 00320.x
Suvorov, R. (2015). The use of visual cues in L2 listening tests. Language Testing, 32(3), 337-357.
Sweller, J. (2011). Cognitive load theory. Psychology of Learning and Motivation, 55, 37-76.
Sydorenko, T. (2010). Modality of input and vocabulary acquisition. Language Learning & 14(2), 50-73.
Taherdoost, H. (2016). Validity and reliability of the research instrument: How to test the validation of a questionnaire/survey in a research. International Journal of Academic Research in Management, 5(3), 28–36.
Tavakol, M., & Dennick, R. (2011). Making sense of Cronbach's alpha. International Journal of 2, 53–55.
Tweed, M., & Wilkinson, T. (2012). Diagnostic testing and educational assessment. The Clinical, 9(5), 299-303.

Vandergrift, L., & Goh, C. C. M. (2012). Teaching and learning second language listening: Metacognition in
Vandergrift, L., Goh, C., Mareschal, C. and Tafaghodtari, M. H. (2006). "The metacognitive awareness listening questionnaire (MALQ): Development and validation". Language Learning, 56 (3), 431-462.
Wagner, E. (2010). "The effect of the use of video texts on ESL listening test-taker performance". Language Testing, 27, 493-513. doi:10.1177/0265532209355668
Wagner, E. (2010). The effect of the use of video texts on ESL listening test-taker performance, 27(4), 493-513.
Warschauer, M. (2003). Technology and social inclusion: Rethinking the digital divide. MIT Press.
Watkins, J., & Wilkins, M. (2011). Using YouTube in the EFL classroom. Language Education in Asia, 2(1), 113-119.