ISSN: 2583-4371

International Journal of Teaching, Learning and Education (IJTLE)

Vol-5, Issue-5, Sep-Oct, 2026

CrossRef DOI: 10.22161/ijtle

IJTLE

Peer-reviewed
Teaching, Learning and Education
IJTLE.COM

editor.ijtle@gmail.com
https://ijtle.com/

International Journal of Teaching, Learning and Education (IJTLE)

ISSN: 2583-4371

DOI: 10.22161/ijtle

Vol-5 Issue-5

September – October 2026

Editor in Chief

Dr. Luisa Maria Arvide Cambra

Copyright © 2026 International Journal of Teaching, Learning and Education

Publisher

International Journal of Teaching, Learning and Education

chiefeditor.ijtle@gmail.com
https://ijtle.com/

About IJTLE

International Journal of Teaching, Learning and Education (IJTLE) is an open-access Peer-Reviewed International journal in teaching, learning, English Language, Humanities, Education Development, Social Science and English Literature.

IJTLE Journal covers but is not limited to the following topics:

Anthropology, applied linguistics studies, Arts, Business Studies, Communication Studies, communicative language teaching (CLT), comparative literature, computational linguistics, Corporate Governance, corpus linguistics, Criminology, critical theory today, Cross-Cultural Studies, cross-cultural studies, Demography, Development Studies, discourse analysis, Economics, Education, English globalization, English language, English language testing and assessment, English literature, English speaking culture, English teaching and learning, Ethics, gender studies in language, Geography, History, Human Rights, Industrial Relations, Information Science, interdisciplinary approaches in literature, International Relations, Law, Library Science, Linguistics, Literature, literature and media, literature studies, Media Studies, methodology, Paralegal, Performing Arts, performing arts (music theater & dance), Philosophy, Political Science, Population Studies, Psychology, Public Administration, Religious Studies, Humanities, second language acquisition, Social Welfare, Sociology, syllabus design and curriculum development, task-based language teaching (tblt), translation studies, Visual Arts, women studies, world literature..etc.

Important Links:

- Author Guidelines
- Peer Review Process
- Publication Policy and Ethics
- Paper Format
- Online Manuscript Submission
- Copyright Form

Email ID: editor.ijtle@gmail.com

International Editorial/ Reviewer Board

- Dr. Andrew Sagayadass Philominraj; Associate Professor, Director, Ph.D. in Education in Consortium, Academic & Researcher, School of English Pedagogy Department of English, Faculty of Education, Catholic University of Maule Talca – Chile

- Dr. Luisa Maria Arvide Cambra; Professor with Chair/ Arabic and Islamic Studies// Humanities Dept. of Philology, University of Almeria. LA Cañada S/N. 04120-Almeria, Spain

- Sarath W. Samaranayake; PhD in Linguistics, English Lecturer, Curriculum and Instruction, Silpakorn University, Sanam Chandra Campus, Nakhon Pathom, Thailand

- Darren Rey C. Javier; Licensed Professional Teacher, Senior High School HUMSS Teacher II, Baras-Pinugay Integrated High School Master of Arts in Education with specialization in English Language Teaching | Philippine Normal University, Philippines

- Almighty C. Tabuena, LPT, MAEd (CAR), PhD (h.c.) PhD in Education (h.c.) Master of Arts in Education, Philippine Normal University, Manila, Philippines

- Dr. Sadia Irshad; Assistant Professor, Thesis Coordinator (M.Phil & PhD Programs), Department of English, Faculty of Social Sciences, Air University, Islamabad, Pakistan.+

- Dr. Ayse DEMIR; School of Foreign Languages, Pamukkale University, Denizli, TURKEY

- Febini M Joseph; Assistant Professor, Department of Basic Science & Humanities, SCMS School of Engineering and Technology, Kerala, India

- Khalid Aada; Lecturer II, Arabic as Critical Language, Culture and Civilization, French linguistics and Literature, Spanish linguistics for Non-Native Speakers, University of Texas, One West University Blvd., Brownsville, Texas

- Jânderson Albino Coswosk; Ph.D. in Literary Studies, Professor of Basic, Technical and Technological Education, Department of Languages, Federal Institute of Espírito Santo,

- Akmaljon Odilov Odilovich
  Head of Department of Tourism Management, Silk Road International University of Tourism and Cultural Heritage

- Dr. Peer Salim Jahangeer
  GDC Kokernag Anantnag J & K (Lecturer); GDC Boys Anantnag J&K (Approved IGNOU Counsellor)

Vol-5, Issue-5; September - October 2026
(10.22161/ijtle.5.5)

On Literary Cadence and Emotional Expression in Prose Translation: A Comparative Analysis of Human and Machine Translation
Authors: Yan Jiayao, Fan Zhixin, Fan Jiying
DOI: 10.22161/ijtle.5.5.1

Page No: 1-11

The Role of Audiovisual Materials in Enhancing English Speaking Skills: A Systematic Review (2020–2026)
Authors: Anwer Mohammed Elhassan Hussein
DOI: 10.22161/ijtle.5.5.2

Page No: 12-20

Transforming Higher Education through Digital Technologies: A Critical Analysis
Authors: Dr. Sheeba Parveen
DOI: 10.22161/ijtle.5.5.3

Page No: 21-23

Contribution of School Management Strategies to Student's Academic Performance in Primary School in Misungwi District, Tanzania
Authors: Christopher John Legonda
DOI: 10.22161/ijtle.5.5.4

Page No: 24-30

AI-Customized Digital Picture Books to Enhance Vocabulary Learning in Early ELT Classrooms: A Quasi-Experimental Study
Authors: Banoth Suresh, Dr. Narendar. N
DOI: 10.22161/ijtle.5.5.5

Page No: 31-40

International Journal of Teaching, Learning and Education (IJTLE)
ISSN: 2583-4371
Vol-5, Issue-5, September-October 2026
Journal Home Page: https://ijtle.com/
Journal DOI: 10.22161/ijtle

On Literary Cadence and Emotional Expression in Prose Translation: A Comparative Analysis of Human and Machine Translation

Yan Jiayao[1], Fan Zhixin[2], Fan Jiying[3]

[1,2]Researcher, Beijing Institute of Petrochemical Technology, China
[3]Associate Professor, Beijing Institute of Petrochemical Technology, China *Corresponding Author and Supervisor

Received: 13 Aug 2026, Received in revised form: 10 Sep 2026, Accepted: 14 Sep 2026, Available online: 19 Sep 2026
©2026 The Author(s). Published by IJTLE. This is an open-access article under the CC BY license (https://creativecommons.org/licenses/by/4.0/).

Fund— This paper is funded by the URT (Undergraduate Research Training) program of the Beijing Institute of Petrochemical Technology, and the number is 2026J00268.

Abstract

Modern Chinese prose takes "loose in form, unified in spirit" as its core trait, its language combining pictorial narrative, rhythmic beauty and restrained lyricism. The central demand of prose translation is therefore not merely the transfer of meaning, but the reproduction of the distinctive layers of literary cadence and the deep-lying emotion of the source text. Artificial-intelligence translation is now easily accessible, yet in the field of literary prose a marked aesthetic gap remains between machine output and the translations of experienced human translators. This paper selects three classic lyric prose works — Zhu Ziqing's "The Sight of My Father's Back" , "Transient Days" and Yu Dafu's "Autumn in the Old Capital" and takes Zhang Peiji's authoritative human translation and the machine translation produced by the AI model DeepSeek as its comparative corpus. Grounded in Nida's theory of functional equivalence, it carries out close reading and comparative analysis along two core dimensions: the reproduction of literary cadence and the expression of emotion. The study finds that Zhang Peiji's human translation attends at once to sentence rhythm, rhetorical texture and stylistic temperament, restoring both the outward cadence of the prose and the implicit emotions concealed beneath scenery, action and narrative. DeepSeek's machine translation, while guaranteeing basic semantic accuracy and grammatical fluency, is generally characterised by flattened syntax, loss of rhetorical beauty, thin emotional layering and an absence of imaginative openness, so that it can hardly achieve the aesthetic functional equivalence that literary translation demands. Drawing on cases from several types of prose, this paper summarises the differences between human and machine translation and proposes a practical pathway for literary prose translation and for human–machine collaborative translation, thereby offering a reference for research on modern Chinese prose in English translation and for the practice of literary translation.

Keywords*— *prose translation; literary cadence; emotional expression; Zhang Peiji; DeepSeek; human–machine translation comparison

I. INTRODUCTION

Artificial-intelligence translation technology is maturing with every iteration. Intelligent translation tools such as DeepL, DeepSeek and GPT are widely used in the translation of all kinds of everyday texts thanks to their outstanding efficiency and low cost.

Unlike ordinary text translation, however, literary translation imposes exceptionally high and specialised requirements on linguistic aesthetics, the creation of imaginative atmosphere and the transmission of emotion — and this is precisely the bottleneck that machine translation has so far been unable to break through. Prose differs in essence from practical genres such as scientific writing and news reporting: it does not pursue straightforward information output, but relies on sentences of alternating length, delicate and varied rhetoric and a lyricism built on blank space to shape a distinctive cadence and emotional atmosphere. Zhang Peiji's *Selected Modern Chinese Essays in English Translation* is a benchmark version in the field of Chinese prose rendered into English; its language is elegant and condensed and fits the spirit of the original works closely, making it a classic model of human literary translation. Setting this classic human translation against DeepSeek's machine output can reveal directly the gap between human and machine translation in handling the cadence and emotion of prose, and therefore possesses ample research value[2].

At the level of theoretical research, existing comparative studies of human and machine prose translation are mostly confined to the analysis of lexical and syntactic correctness, and their research dimension is relatively narrow. Focusing on the core aesthetic elements of prose translation and taking literary cadence and emotion as its point of entry, this paper further refines the evaluative system and criteria of literary translation and effectively broadens and enriches the research perspective and content of studies on modern Chinese prose in English translation. At the level of practical application, by systematically comparing the merits and shortcomings of human literary translation and AI machine translation, this paper clarifies the inherent limitations of machine translation in the field of literary translation and explores human–machine collaborative approaches suited to prose translation, so as to provide realistic and workable reference for the translation practice of students of literary translation and for their polishing and optimisation of translations.

II. LITERATURE REVIEW

2.1 Definition of Core Concepts

Literary cadence (文韵) refers to the aesthetic quality of language produced by the rhythm of composition, the cadence of sentence patterns, the texture of rhetoric, stylistic temperament and sensory description. It is the central manifestation of the outward literary beauty of prose and an important feature distinguishing prose from practical texts. Emotional expression encompasses both the explicit lyricism stated directly in a text and the implicit emotion concealed in scenery, action and narrative, covering multi-layered states of mind such as kinship, melancholy, loneliness and attachment; it is the key to the literary value and affective power of prose.

The choice of literary cadence and emotional expression as the core evaluative concepts of this study follows from two considerations: the nature of the text itself and the current state of research. As far as prose is concerned, practical texts take the transfer of information as their primary goal and are judged chiefly on accuracy of information and clarity of logic, whereas prose writing is not limited to the statement of events and facts: on the one hand it shapes a distinctive cadence of language by arranging sentence patterns and the polishing rhetoric, and on the other it moves the reader through the expression of emotion at several levels. Literary cadence and emotion correspond precisely to the “form” and the “spirit” of prose, so that grasping these two dimensions is to grasp the most central yardstick for evaluating Chinese prose in English translation.

2.2 Theoretical Framework

This paper is grounded in Nida's theory of functional equivalence. According to Nida, a translation should achieve equivalence on four levels — lexical, syntactic, textual, and stylistic — and the ultimate criterion for translation quality is that the target-language reader should enjoy a reading experience identical to that of the source-language reader. When this framework is applied to prose translation, a well-executed human translation is capable of attaining functional equivalence in three dimensions: meaning, style, and emotion, thereby reproducing the tone and atmosphere of the original. Machine translation, by contrast, tends to achieve only surface-level equivalence at the lexical and syntactic levels, while falling short at the higher levels of stylistic beauty and emotional resonance. As a result, it can scarcely satisfy the demands of literary prose translation[7].

2.3 Current Research at Home and Abroad

Scholarship on the English translation of prose and on the comparison of intelligent human and machine translation has already accumulated a certain number of results. In the study of human prose translation, existing research has focused largely on the translation

Crossref DOI: https://dx.doi.org/10.22161/ijtle.5.5.1

style and aesthetic reproduction of Zhang Peiji's prose translations, generally acknowledging that his versions are elegant and expressive and reproduce the spirit of the originals to a high degree, making them benchmark versions of modern prose in English [2]. As for the three classic essays "The Sight of My Father's Back", "Transient Days" and "Autumn in the Old Capital", previous research has largely focused on the translation techniques of the human version, its handling of lexis and syntax, and the analysis of artistic mood in a single text. At the level of translation technique and stylistic reproduction, Yu Yuanling used "The Sight of My Father's Back" as a case to demonstrate the translatability of prose style and analysed the specific techniques by which Zhang Peiji's version reproduces the style of the original [14]. At the level of lexical and syntactic treatment, drawing on a quantitative, corpus-based approach, Wang Weihua and Lei Pengfei examined the lexical and syntactic features of the English translation of "Transient Days"[15], while Liu Rui compared the differences in rhythm and syntax between Zhang Peiji's and Zhu Chunshen's versions of the same essay from the perspective of literary stylistics[16]. Jiang Hui and Yu Miao appreciated the art of diction in Zhang Peiji's translation of "Autumn in the Old Capital" against Liu Miqing's aesthetic criteria of "accuracy, beauty and refinement"[17]. At the level of artistic-mood analysis, He Jiandi and Tang Lijun explored, in terms of reception aesthetics, the strategies by which the translation of "Autumn in the Old Capital" conveys its emotional tone and artistic conception [18], while Pan Xing analysed the reproduction of beauty in sound, form and meaning in that version on the basis of Xu Yuanchong's theory of the "three beauties". [19] These studies, however, either focus on translation technique itself or are limited to the case analysis of a single text, and their perspectives remain fairly traditional and one-dimensional.

In the study of machine translation and human–machine translation comparison, existing results mostly take practical texts such as news and scientific writing as their object, and specialised comparative studies devoted to the prose genre are relatively rare. The comparative studies of human and machine prose translation that do exist are mostly confined to surface-level linguistic analysis such as lexical correctness and syntactic normativity, and very few of them focus on the cadence, compositional rhythm and implicit emotional expression that lie at the core of prose. Although Guo Wanyu and Wang Honghua used "The Sight of My Father's Back" as their corpus to compare the differences in emotional expression between Zhang Peiji's human translation and a large-language-model translation [20], they addressed only the single dimension of emotion and did not undertake a systematic examination at the level of literary cadence. It is generally acknowledged in the field that although AI translation is efficient and accurate in its semantic output, it suffers from stiff syntax, absent rhetoric and weakened artistic mood, and can hardly achieve the aesthetic and emotional equivalence required of literary translation.

Overall, a systematic comparative analysis of benchmark human translations and machine translation organised around the two core dimensions of literary cadence and emotional expression offers an innovative point of entry and considerable room for research.

III. RESEARCH DESIGN AND CORPUS

3.1 Objects of Comparison

This paper selects two kinds of translation as its comparative corpus. For the human translation it adopts the standard version in Zhang Peiji's *Selected Modern Chinese Essays in English Translation*, a version widely recognised in Chinese academia as the authoritative model for rendering modern prose into English [2]; for the machine translation it uses the standard output generated by DeepSeek, currently a mainstream general-purpose artificial-intelligence translation model.

3.2 Corpus Selection and Classification

This paper selects three classic modern prose works — "The Sight of My Father's Back", "Transient Days" and "Autumn in the Old Capital" — as its research corpus, with the following specific passages: from "The Sight of My Father's Back", the classic episode in which the father climbs over the railway tracks to buy oranges for "me"; from "Transient Days", the opening landscape parallelism built on swallows, willows and peach blossoms, together with the passage of self-reflection at the end; and from "Autumn in the Old Capital", three passages — the contrast between northern and southern autumn, the description of the autumn scene in the courtyard, and the closing expression of feeling. All the passages selected are the core fragments of each essay that display the greatest richness of cadence and the highest concentration of emotion, fully embodying the literary qualities of prose in sentence rhythm,

Crossref DOI: https://dx.doi.org/10.22161/ijtle.5.5.1

rhetorical devices and emotional expression, and providing a representative corpus for the subsequent comparative analysis of human and machine translation.

3.3 Dimensions of Comparative Analysis

This paper carries out its comparative analysis of the translation along two dimensions. The first is the reproduction of literary cadence, focusing on the rhythmic effect of parallel structures, the compositional rhythm created by the alternation of long and short sentences, the textual texture produced by various rhetorical devices and sensory description, and the overall stylistic temperament of the text. The second is emotional expression, analysing chiefly the explicit and direct lyricism in the translations, the implicit emotion concealed in the description of scenery and action, and the degree to which the overall atmosphere and multi-layered emotions of the original are reproduced.

3.4 Research Content and Methods

This paper takes three pieces of classic modern prose works of different kinds — "The Sight of My Father's Back", "Transient Days" and "Autumn in the Old Capital" — as its corpus, covering respectively the narrative of family affection, philosophical lyricism and landscape lyricism. By comparing the specific differences between Zhang Peiji's classic human translation and DeepSeek's machine translation in the reproduction of literary cadence and the expression of emotion, it explores in depth the internal causes of the differences between human and machine translation and, on that basis, summarises an optimised pathway suited to the English translation of modern prose. The overall research methodology combines documentary research, comparative textual analysis, and empirical case analysis.

IV. A HUMAN–MACHINE COMPARATIVE ANALYSIS OF THE EXPRESSION OF LITERARY CADENCE IN PROSE TRANSLATION

Each of the following examples consisted five elements including the source text in Chinese; the literal rendering (baseline) which is provided not as a competing translation but as a baseline against which the stylistic choices of the human translator and the machine translation system can be measured; Zhang Peiji's translation, representing the human translation in this study, which is cited from *Selected Modern Chinese Essays in English Translation* published by Shanghai Foreign Language Education Press in 2007; DeepSeek's translation version which is AI translation; and a comparative analysis of the two versions.

4.1 Cadence of the Parallel Structures Opening in "Transient Days"

Source Text (Chinese)

> 燕子去了，有再来的时候；杨柳枯了，有再青的时候；桃花谢了，有再开的时候。但是，聪明的，你告诉我，我们的日子为什么一去不复返呢？—— 是有人偷了他们罢：那是谁？又藏在何处呢？是他们自己逃走了罢：现在又到了那里呢？

> *Literal rendering(Baseline):* Swallows have gone, but there is a time when they come back; willows have withered, but there is a time when they turn green again; peach blossoms have faded, but there is a time when they bloom again. But tell me, you who are wise, why do our days go by and never return? — Perhaps someone has stolen them: who is it? And where are they hidden? Perhaps they have run away of their own accord: and where are they now?

Zhang Peiji's Human Translation

> If swallows go away, they will come back again. If willows wither, they will turn green again. If peaches shed their blossoms, they will flower again. But, tell me, you the wise, why should our days go by never to return? Perhaps they have been stolen by someone. But who could it be and where could he hide them? Perhaps they have just run away by themselves. But where could they be at the present moment?

DeepSeek's Machine Translation

> Swallows may depart, yet they return in time; willows may wither, yet they green again; peach blossoms may fall, yet they bloom afresh. But, O wise one, tell me—why should our days be gone never to come back? Could it be that someone has stolen them away? But who, and where are they hidden? Or have they fled on their own? And where, then, have they now arrived?

Comparative Analysis of Literary Cadence

The three sets of parallel short sentences in the source

Crossref DOI: https://dx.doi.org/10.22161/ijtle.5.5.1

text are neatly symmetrical, and their recurrence and return create a slow, lingering rhythm that is the signature cadence of "Transient Days". Zhang Peiji reproduces the parallel structure with a uniform "If..." conditional clause, balancing the length of his sentences so that the rhythm fits the cyclical beauty of the original; the syntax is orderly and gentle and preserves the warm, mellow verbal cadence of the prose.

DeepSeek retains the parallel structure, but the sentences are joined stiffly, with heavy use of the adversative conjunction "yet" splitting the lines apart and breaking the smooth, unhurried rhythm of the original when read aloud; its word choice is bookish and stiff, losing the natural, plain prose style of the original, and its rhythmic layering is weaker than that of the human version. The human translator actively reshapes English syntax to match the sentence rhythm of the original, whereas the machine merely replicates the literal parallelism mechanically and cannot attend to the beauty of the rhythm when read aloud.

4.2 Cadence of Narrative-Action Rhythm in the Orange-Buying Scene of "The Sight of My Father's Back"

Source Text (Chinese)

> 我看见他戴着黑布小帽，穿着黑布大马褂，深青布棉袍，蹒跚地走到铁道边，慢慢探下身去，尚不大难。可是他穿过铁道，要爬上那边月台，就不容易了。他用两手攀着上面，两脚再向上缩；他肥胖的身子向左微倾，显出努力的样子。这时我看见他的背影，我的眼泪很快地流下来了。
>
> *Literal rendering(Baseline):* I saw him wearing a small black cloth cap, a black cloth mandarin jacket and a dark blue cotton-padded gown, hobbling to the edge of the railway track and slowly stooping down, which was not yet too difficult. But for him to cross the track and climb up onto the platform on the other side was not easy. He clutched the top with both hands and drew his feet up; his stout body leaned slightly to the left, showing the effort he was making. At this moment I saw his back, and my tears quickly streamed down.

Zhang Peiji's Human Translation

> I watched him hobble towards the railway track in his black skullcap, black cloth mandarin jacket and dark blue cotton-padded cloth long gown. He had little trouble climbing down the railway track, but it was a lot more difficult for him to climb up that platform after crossing the railway track. His hands held onto the upper part of the platform, his legs huddled up and his corpulent body tipped slightly towards the left, obviously making an enormous exertion. While I was watching him from behind, tears gushed from my eyes.

DeepSeek's Machine Translation

> I watched him, wearing a small black cloth cap, a black mandarin jacket, and a dark blue cotton-padded gown, as he hobbled to the edge of the railway track. He bent down slowly to step onto the tracks—that was not too difficult. But crossing over and climbing up the opposite platform was no easy feat for him. He grasped the upper edge with both hands, then drew his feet up; his stout body leaned slightly to the left, straining with effort. At that moment, as I gazed at his retreating figure, tears welled up and streamed down my cheeks.

Comparative Analysis of Literary Cadence

The continuous action description of the source text alternates long and short sentences, its rhythm slow and heavy, heightening the pictorial sense of the father's age and clumsiness. Zhang Peiji selects graded verbs such as "hobble", "huddled up" and "enormous exertion"; his sentences are neither too loose nor too tight — the long sentence lays out the whole figure while the short sentences highlight the effort of climbing, so that the rhythm of the writing rises and falls with the action and the slow, oppressive narrative cadence of the original is restored.

DeepSeek's version chops the syntax into fine fragments, and a large number of short sentences and dashes break the fluency of the narrative; its verbs, such as "grasp" and "straining with effort", are plain and monotonous and do not distinguish between the weight of different actions. The rhythm is uniform throughout, without rise or fall, dissolving the pictorial rhythm inherent in the original action description; the writing is plain and lacks layering, and its literary cadence is thin.

Crossref DOI: https://dx.doi.org/10.22161/ijtle.5.5.1

4.3 Stylistic Texture of the Courtyard Landscape Description in "Autumn in the Old Capital"

Source Text (Chinese)

在北平即使不出门去罢，就是在皇城人海之中，租人家一椽破屋来住着，早晨起来，泡一碗浓茶，向院子一坐，你也能看得到很高很高的碧绿的天色，听得到青天下驯鸽的飞声。从槐树叶底，朝东细数着一丝一丝漏下来的日光，或在破壁腰中，静对着像喇叭似的牵牛花（朝荣）的蓝朵，自然而然地也能够感觉到十分的秋意。

Literal rendering(Baseline): In Beiping, even if you do not go out, living in a rented broken-raftered house amid the sea of people in the imperial city, when you rise in the morning, brew a bowl of strong tea and sit down in the courtyard, you can still see the very high, very high azure sky and hear the flight of the tame pigeons under the clear heavens. Counting strand by strand the sunlight that leaks down from beneath the leaves of the locust trees towards the east, or quietly facing the blue trumpet-shaped blossoms of the morning glory on the broken wall, you will naturally and spontaneously feel autumn to the full.

Zhang Peiji's Human Translation

Suppose you put up in a humble rented house inside the bustling imperial city, you can, on getting up at dawn, sit in your courtyard sipping a cup of strong tea, leisurely watch the high azure skies and listen to pigeons circling overhead. Turn eastward under locust trees to closely observe streaks of sunlight filtering through their foliage, or quietly watch the trumpet-shaped blue flowers of morning glories climbing half way up a dilapidated wall, and an intense feeling of autumn will of itself well up inside you.

DeepSeek's Machine Translation

Even if you never set foot outside the door in Peiping—living in a ramshackle rented room amid the vast sea of humanity within the imperial city—you can still, upon rising in the morning, brew a bowl of strong tea, sit in the courtyard, and gaze up at the immensely high, azure sky, while listening to the sound of tame pigeons circling overhead. Look eastward from beneath the locust trees and count, one by one, the thin threads of sunlight filtering through the leaves; or quietly fix your gaze upon the blue, trumpet-shaped blooms of morning glories climbing along a broken wall—and before you know it, a full measure of autumn feeling will arise quite naturally within you.

Comparative Analysis of Literary Cadence

Yu Dafu's original is cool, light and elegantly restrained in style, creating a relaxed and quiet atmosphere out of small, ordinary scenes. Zhang Peiji selects refined and controlled vocabulary such as "humble", "azure" and "dilapidated", and differentiates states of action with distinct verbs such as "sipping", "leisurely watch" and "quietly watch"; his sentences are unhurried and leave space, fitting the light, solitary style of the original.

DeepSeek's version piles up vocabulary, with words such as "ramshackle" and "immensely high" exaggerated and blunt, and its many dashes and semicolons destroy the relaxed, unhurried rhythm of the prose. The writing is heavy and congested, losing the simple, cool and reserved cadence that is unique to Yu Dafu's prose, and the aesthetic texture of the landscape writing drops sharply.

4.4 Summary

Human translation can adjust sentence rhythm, differentiate levels of verbs and match the stylistic texture of the original according to the generic features of prose, preserving intact the rhythm and verbal cadence unique to prose; DeepSeek's machine translation completes only the literal transfer of information, with fragmented syntax, homogeneous vocabulary and poor stylistic fit, and is generally marked by a flattening of literary cadence and a loss of aesthetic beauty[7][9].

V. A HUMAN–MACHINE COMPARATIVE ANALYSIS OF EMOTIONAL EXPRESSION IN PROSE TRANSLATION

5.1 The Orange-Buying Scene in "The Sight of My Father's Back": Restoration of Implicit Family Affection

Source Text (Chinese)

我看见他戴着黑布小帽，穿着黑布大马褂，深青布棉袍，蹒跚地走到铁道边，慢慢探下身去，尚不大难。可是他穿过铁道，要爬上那边月台，就不容易了。他用两手攀着上面

Crossref DOI: https://dx.doi.org/10.22161/ijtle.5.5.1

，两脚再向上缩；他肥胖的身子向左微倾，显出努力的样子。这时我看见他的背影，我的眼泪很快地流下来了。

Literal rendering(Baseline): I saw him wearing a small black cloth cap, a black cloth mandarin jacket and a dark blue cotton-padded gown, hobbling to the edge of the railway track and slowly stooping down, which was not yet too difficult. But for him to cross the track and climb up onto the platform on the other side was not easy. He clutched the top with both hands and drew his feet up; his stout body leaned slightly to the left, showing the effort he was making. At this moment I saw his back, and my tears quickly streamed down.

Zhang Peiji's Human Translation

I watched him hobble towards the railway track in his black skullcap, black cloth mandarin jacket and dark blue cotton-padded cloth long gown. He had little trouble climbing down the railway track, but it was a lot more difficult for him to climb up that platform after crossing the railway track. His hands held onto the upper part of the platform, his legs huddled up and his corpulent body tipped slightly towards the left, obviously making an enormous exertion. While I was watching him from behind, tears gushed from my eyes.

DeepSeek's Machine Translation

I watched him, wearing a small black cloth cap, a black mandarin jacket, and a dark blue cotton-padded gown, as he hobbled to the edge of the railway track. He bent down slowly to step onto the tracks—that was not too difficult. But crossing over and climbing up the opposite platform was no easy feat for him. He grasped the upper edge with both hands, then drew his feet up; his stout body leaned slightly to the left, straining with effort. At that moment, as I gazed at his retreating figure, tears welled up and streamed down my cheeks.

Comparative Analysis of Emotion

Not a single sentence in this passage expresses feeling directly; all the fatherly love, guilt and heartache are embedded in the series of details of the father's climbing, making it a typical passage of implicit lyricism. Zhang Peiji precisely selects words such as "hobble" (walking unsteadily), "corpulent" (bulky and aged) and "enormous exertion" (straining to the utmost) to portray the father's aged and laborious bearing, and the closing word "gushed" vividly captures the surging burst of tears, giving concrete form to the intense guilt and gratitude buried in the narrator's heart and building up, layer by layer, the reserved and deep affection between father and son; the emotional layering is delicate and full.

DeepSeek's machine translation, by contrast, merely renders the surface actions in a flat, literal way; words such as "stout" and "straining with effort" are plain and weaken the image of the father's aged helplessness, while "tears well up" is only an objective description of the act of crying and lacks the explosive force of feeling, so that it cannot convey the restrained and weighty guilt behind the words: the expression of family affection is thin and weak and can hardly move the reader.

5.2 The Closing Introspection of "Transient Days": Melancholy over Time and Self-Reflection

Source Text (Chinese)

过去的日子如轻烟，被微风吹散了，如薄雾，被初阳蒸融了；我留着些什么痕迹呢？我何曾留着像游丝样的痕迹呢？我赤裸裸来到这世界，转眼间也将赤裸裸的回去罢？但不能平的，为什么偏要白白走这一遭啊？你聪明的，告诉我，我们的日子为什么一去不复返呢？

Literal rendering(Baseline): The days past are like light smoke, dispersed by the breeze, like thin mist, melted away by the early sun; what traces have I left behind? Have I ever left a trace like a gossamer thread? I came into this world stark naked, and in the twinkling of an eye I shall go back stark naked too? But what I cannot accept is this: why must I pass through this once for nothing at all? You who are wise, tell me, why do our days go by and never return?

Zhang Peiji's Human Translation

The bygone days, like wisps of smoke, have been dispersed by gentle winds, and, like thin mists, have been evaporated by the rising sun. What traces have I left behind? No, nothing, not even gossamer-like traces. I have come to this world stark naked, and in the twinkling of an eye, I am to go back as stark

Crossref DOI: https://dx.doi.org/10.22161/ijtle.5.5.1

naked as ever. However, I am taking it very much to heart: why should I be made to pass through this world for nothing at all?

O you the wise, would you tell me please: why should our days go by never to return?

DeepSeek's Machine Translation

The days that are past—like wisps of smoke, they have been scattered by the light breeze; like thin mists, they have been dissolved by the morning sun. What trace have I left behind? Have I ever left so much as a gossamer thread? Naked I came into this world, and in the blink of an eye, naked I shall return. Yet I cannot reconcile myself to this—why should I have to make this journey through life for nothing at all?

O wise one, tell me—why should our days be gone never to come back?

Comparative Analysis of Emotion

The source text uses the two metaphors of light smoke and thin mist together with a series of self-questionings to express the unwillingness of time wasted and a spiritual self-anxiety; the emotion is restrained and deep, hidden beneath the surface of the words. Zhang Peiji uses "gossamer-like" to restore the elusive, insubstantial image of time, while "take it very much to heart" precisely conveys the regret of an unresolved knot in the heart, and the closing question preserves the melancholy of the original's repeated probing, fully reproducing the author's complex mood of self-examination and pervasive sorrow.

DeepSeek merely translates the metaphors and questions mechanically and literally, with no modifying vocabulary to heighten the emotion: it objectively states the surface facts of time passing and life being void, and cannot convey the tangled, unwilling self-reflection deep within the words. The intense sentimental atmosphere of the original is largely dissolved.

5.3 The Opening of "Autumn in the Old Capital": The Wanderer's Loneliness and Attachment to His Native Place

Source Text (Chinese)

秋天，无论在什么地方的秋天，总是好的；可是啊，北国的秋，却特别地来得清，来得静，来得悲凉。我的不远千里，要从杭州赶上青岛，更要从青岛赶上北平来的理由，也不过想饱尝一尝这"秋"，这故都的秋味。江南，秋当然也是有的；但草木凋得慢，空气来得润，天的颜色显得淡，并且又时常多雨而少风；一个人夹在苏州上海杭州，或厦门香港广州的市民中间，浑浑沌沌地过去，只能感到一点点清凉，秋的味，秋的色，秋的意境与姿态，总看不饱，尝不透，赏玩不到十足。

Literal rendering(Baseline): Autumn, autumn wherever it is, is always good; and yet autumn in the North is especially clear, especially quiet, especially desolate. The reason I have not counted a thousand miles far in hurrying from Hangzhou to Qingdao, and then from Qingdao to Beiping, is no more than a wish to savour this "autumn" to the full, the autumn flavour of the old capital. In Jiangnan autumn of course also exists; but there the plants wither slowly, the air is moist, the colour of the sky is pale, and it is often rainy and rarely windy; caught among the townspeople of Suzhou, Shanghai, Hangzhou, or Xiamen, Hong Kong and Guangzhou, one drifts along in a daze and can feel only a faint touch of coolness — the flavour of autumn, the colour of autumn, the mood and posture of autumn can never be seen to the full, tasted through, or enjoyed completely.

Zhang Peiji's Human Translation

Autumn, wherever it is, always has something to recommend itself. In North China, however, it is particularly limpid, serene and melancholy. To enjoy its atmosphere to the full in the onetime capital, I have, therefore, made light of travelling a long distance from Hangzhou to Qingdao, and thence to Peiping.

There is of course autumn in the South too, but over there plants wither slowly, the air is moist, the sky pallid, and it is more often rainy than windy. While muddling along all by myself among the urban dwellers of Suzhou, Shanghai, Xiamen, Hong Kong or Guangzhou, I feel nothing but a little chill in the air, without ever relishing to my heart's content the flavour, colour, mood and style of the season.

DeepSeek's Machine Translation

Autumn—wherever it is found—is always a

Crossref DOI: https://dx.doi.org/10.22161/ijtle.5.5.1

> season to be cherished. Yet autumn in the North comes upon one with a peculiar clarity, a stillness, a melancholy all its own. The reason I have traveled all the way from Hangzhou to Qingdao, and then from Qingdao on to Peiping, is simply to drink in this very "autumn"—the very essence of autumn in the former capital.
>
> To be sure, autumn exists in the South as well. But there the plants wither more slowly, the air is more humid, the sky paler, and the weather more rainy than windy. Drifting along in a daze among the city crowds of Suzhou, Shanghai, Hangzhou, or for that matter Xiamen, Hong Kong, one feels only a faint touch of coolness—never quite feasting one's eyes on the hues of autumn, never quite savouring its flavours, never quite absorbing its moods and postures to the full.

Comparative Analysis of Emotion

The three parallel phrases "clear, quiet and desolate" establish the solitary undertone of the whole essay, and the journey of a thousand miles to Beiping to enjoy autumn conceals the particular obsession and homesickness of a wandering son. Zhang Peiji chooses the three progressively deepening adjectives "limpid, serene and melancholy", precisely delineating the cool sadness unique to northern autumn; "make light of travelling" highlights the author's obstinate insistence on reaching the old capital whatever the distance, fully restoring the wanderer's complex state of mind — his loneliness and his attachment to his native place.

DeepSeek's version makes no distinction of level in its vocabulary and simply translates "clear, quiet, sad", while stating the act of travelling flatly; it cannot convey the deep feeling that leads the author to cross a thousand miles for a single taste of autumn, and stays only at the surface description of scenery, losing the loneliness and obsession of the wanderer that underlie the words.

5.4 The Ending of "Autumn in the Old Capital": Intense Lyricism of Extreme Attachment

Source Text (Chinese)

> 秋天，这北国的秋天，若留得住的话，我愿把寿命的三分之二折去，换得一个三分之一的零头。
>
> *Literal rendering(Baseline):* Autumn, this autumn of the North, if it could be kept, I would be willing to cut away two-thirds of my life and exchange it for the remaining one-third.

Zhang Peiji's Human Translation

> Autumn, I mean Northern autumn, if only it could be made to last forever! I would be more than willing to keep but one-third of my life-span and have two-thirds of it bartered for the prolonged stay of the season!

DeepSeek's Machine Translation

> Autumn—this autumn of the North—if only it could be kept from passing! I would willingly trade away two-thirds of my allotted years, if only to keep the remaining one-third in an eternal autumn.

Comparative Analysis of Emotion

This sentence is the emotional climax of the whole essay, directly expressing the ardent depth of feeling that would give up more than half of one's lifespan to keep the autumn of the North, a love tinged with a faint desolation. Zhang Peiji uses an exclamatory sentence pattern to strengthen the inner expectation, with "be more than willing" and "bartered" highlighting the intense emotion of willing sacrifice; the exclamatory tone fits the ardent yet sorrowful lyrical key of the original.

DeepSeek's version is flat in syntax, and the phrase "trade away" is cold and objective, weakening the tension of subjective feeling: it merely states objectively the act of exchanging one's lifespan and can hardly convey the author's almost obsessive, deep attachment, so that the intense lyricism is greatly diluted.

5.5 Summary

Human translators possess literary empathy and can precisely capture both explicit lyricism and the implicit emotions hidden in action, scenery and blank space, heightening the emotional atmosphere at several levels through vocabulary, syntax and tone; machine translation can recognise only the surface emotion words of a text and cannot interpret the deeper state of mind that lies beyond the words. Its emotional expression is single and flat, and it loses the core charm of the restrained lyricism of prose[6][13].

Crossref DOI: https://dx.doi.org/10.22161/ijtle.5.5.1

VI. CAUSES OF THE DIFFERENCES BETWEEN HUMAN AND MACHINE TRANSLATION AND THEIR IMPLICATIONS FOR PROSE TRANSLATION

6.1 Internal Causes of the Differences between Human and Machine Translation

The internal causes of the differences between human and machine translation lie mainly in the following aspects. First, the underlying operating logic differs. DeepSeek's machine translation relies on probability matching over large data sets and recognises only linguistic symbols and literal meaning; it has no capacity for literary aesthetics or human empathy and cannot perceive sentence rhythm, the blank space of imagery or stylistic mood. Zhang Peiji, as a professional literary translator, possesses a deep literary accumulation, can read the cadence and state of mind behind the words, and actively adjusts syntax and vocabulary to restore the aesthetics and emotion of the original[8][10].

Second, the translation goals differ. The core goal of AI translation is to ensure semantic accuracy and grammatical fluency and to complete the transfer of information, whereas the core goal of human literary translation is to achieve Nida's functional equivalence and to take account of the multiple literary values of meaning, style, emotion and artistic conception.[1][10] Finally, the two differ in text-processing capacity. Machines excel at standardised practical texts that make no aesthetic demands; prose, by contrast, is a highly implicit, highly subjective and highly aesthetic kind of text in which a great deal of emotion lies beyond the words and depends on the translator's subjective aesthetic interpretation — a natural weak point of AI translation[11].

6.2 Practical Implications for Literary Prose Translation

Comparing Zhang Peiji's human translation with DeepSeek's machine output has yielded a number of profound insights into prose translation. On the one hand, in the translation of purely literary prose, human translation is irreplaceable. For classic prose that relies on literary cadence and restrained lyricism, relying on AI output alone will cause a double loss of beauty and emotion, and the aesthetic interpretation and literary polishing performed by a human translator cannot be replaced by a machine. On the other hand, artificial-intelligence translation can serve as an auxiliary tool, making it possible to build a model of human–machine collaborative translation [11][12]. AI is suited to rapidly generating a basic draft and completing the basic semantic conversion, while the human translator can adjust sentence rhythm, improve rhetorical texture, supply implicit emotion and restore the style of the original on that basis, balancing efficiency with literary aesthetics. Finally, the evaluation of Chinese prose in English translation needs to establish a dual standard of "literary cadence + emotion". Judging the quality of a prose translation cannot be a matter of checking lexical and grammatical correctness alone; two aesthetic dimensions — the rhythm of the writing, the fit of the style, and the restoration of implicit emotion — must be added, so as to be closer to the essence of literary translation[9].

VII. CONCLUSION AND RESEARCH OUTLOOK

7.1 Conclusion

Taking three typical modern prose works — "The Sight of My Father's Back", "Transient Days" and "Autumn in the Old Capital" — as its corpus, this paper compares Zhang Peiji's human translation with DeepSeek's machine translation and draws conclusions along the two dimensions of literary cadence and emotion.

First, at the level of literary cadence, human translation can precisely reproduce the parallel rhythms of prose, the narrative rhythm of long and short sentences, the stylistic quality of landscape writing and the texture of rhetoric, preserving intact the aesthetic cadence of the prose; DeepSeek's machine translation has rigid syntax, homogeneous vocabulary and a fractured style, and generally suffers from a loss of literary cadence[7].

Second, at the level of emotion, human translators can capture both explicit lyricism and the implicit emotions hidden in scenery and action, building a multi-layered emotional atmosphere through vocabulary, syntax and tone; machine translation can interpret only literal emotion and cannot handle the restrained lyricism of blank space, so that its emotional expression is flat and thin[6][13].

Third, at the level of causes, the gap between human and machine translation stems from differences in operating logic, translation goals and text-interpretation capacity; AI can complete only the transfer of information, whereas human translators possess literary aesthetics and empathy and can satisfy the multiple aesthetic demands of prose translation.

7.2 Limitations

The corpus of this paper is limited to three modern lyric essays, and the range of subject matter covered is

Crossref DOI: https://dx.doi.org/10.22161/ijtle.5.5.1

narrow; only one AI model, DeepSeek, was used, so that no horizontal comparison of several translation tools could be made. In future work the study could be extended to prose genres such as familiar essays and travel writing, and a wider range of AI models could be included for comparative research.

7.3 Outlook

Artificial-intelligence translation technology will continue to iterate, and the fluency of machine output will continue to improve; literary aesthetics, human empathy and the creation of artistic conception, however, will remain the core strengths of human translation. Human–machine collaborative translation will become the mainstream model for literary prose translation [8][11]: in this model the machine undertakes the basic translation work while the human translator concentrates on polishing literary cadence and restoring emotional and imaginative atmosphere, balancing translation efficiency against literary aesthetic value.

REFERENCES

[1] Nida, E. A., & Taber, C. R. (2004). *The Theory and Practice of Translation* [M]. Shanghai: Shanghai Foreign Language Education Press.

[2] Zhang Peiji. (2007). *Selected Modern Chinese Essays in English Translation, Volume One* [M]. Shanghai: Shanghai Foreign Language Education Press.

[3] Zhu Ziqing. (2020). *Prose of Zhu Ziqing* [M]. Beijing: People's Literature Publishing House.

[4] Yu Dafu. (2019). *Prose of Yu Dafu* [M]. Hangzhou: Zhejiang Literature and Art Publishing House.

[5] Liu Chongde. (1991). *Ten Lectures on Literary Translation* [M]. Beijing: China Translation & Publishing Corporation.

[6] Wu Jianlan, Zhang Luo, & Zhu Hanghui. (2018). Differences between machine translation and human translation: A case study of literary texts [J]. *Overseas English*, (16).

[7] Liu Zhuan, & Tan Ping. (2026). A study of the application of machine translation to prose from the perspective of functional equivalence theory: The case of Zhu Ziqing's "Transient Days" [J]. *Modern Linguistics*, 14(1), 617–623.

[8] Wang Hongtao. (2023). Defending the last bastion: A sociological approach to the challenged literary translation [J]. *Babel*, 69(4).

[9] Hu Yanli. (2023). Prose translation strategy study from Chinese to English from the perspective of functional equivalence [J]. *International Journal of Education and Humanities,* 8(1).

[10] Zhou Lingshun. (2025). From "translationese" to "balanced beauty": The core difference between AI and human translators [N]. *Chinese Social Sciences Today*.

[11] Kolb, W. (2023). Human and machine translation of occasionalisms in literary texts [J]. *Target*, 35(4).

[12] Che Mingming, & Li Hongyu. (2025). A corpus-based comparison of ChatGPT translation and human translation: The case of modern and contemporary Chinese prose [J]. *Translation Research and Teaching*, (3), 104–110.

[13] Fan Wuqiu, & Zhong Hanchun. (2023). On the bottlenecks in machine translation from the perspective of the diversity of thinking modes [J]. *Foreign Language Teaching*, 44(1).

[14] Yu Yuanling. (2004). On the translatability of style seen from the reproduction of "The Sight of My Father's Back": Reflections on reading Zhang Peiji's English translation of the essay [J]. *Journal of Yibin University*, (03).

[15] Wang Weihua, & Lei Pengfei. (2023). A corpus-based study of translator style: The English translation of "Transient Days" as a case [J]. *Journal of Anhui University of Technology (Social Sciences Edition)*, 40(2).

[16] Liu Rui. (2023). A comparative analysis of English translations of "Transient Days" from the perspective of literary stylistics: The cases of Zhang Peiji and Zhu Chunshen [J]. *Comparative Study of Cultural Innovation*, (33).

[17] Jiang Hui, & Yu Miao. (2019). Accuracy, beauty and refinement: An appreciation of diction in English prose translation from the perspective of translation aesthetics — The case of Zhang Peiji's English translation of "Autumn in the Old Capital" [J]. *Wuling Journal*, (06).

[18] He Jiandi, & Tang Lijun. (2023). A study of English prose translation from the perspective of reception aesthetics: The case of Zhang Peiji's English translation of "Autumn in the Old Capital" [J]. *Overseas English*, (15).

[19] Pan Xing. (2023). An appreciation of Zhang Peiji's prose translation from the perspective of translation aesthetics: The case of the English translation of "Autumn in the Old Capital" [J]. *Writers and Masterpieces*, (10).

[20] Guo Wanyu, & Wang Honghua. (2025). A comparative study of emotional expression in human and machine translations of "The Sight of My Father's Back" [J]. *China Terminology*, 27(4), 91–93.

Crossref DOI: https://dx.doi.org/10.22161/ijtle.5.5.1



International Journal of Teaching, Learning and Education (IJTLE)
ISSN: 2583-4371
Vol-5, Issue-5, September-October 2026
Journal Home Page: https://ijtle.com/
Journal DOI: 10.22161/ijtle

The Role of Audiovisual Materials in Enhancing English Speaking Skills: A Systematic Review (2020–2026)

Anwer Mohammed Elhassan Hussein

Research Scholar, Faculty of Graduate Studies, Alzaiem Alazhari University, Sudan
Email: anwermohamad79@gmail.com

Received: 25 Aug 2026, Received in revised form: 21 Sep 2026, Accepted: 28 Sep 2026, Available online: 04 Oct 2026
©2026 The Author(s). Published by IJTLE. This is an open-access article under the CC BY license (https://creativecommons.org/licenses/by/4.0/).

Abstract

Speaking is among the hardest skills for English learners to develop, especially where authentic oral interaction is scarce. Audiovisual (AV) materials such as films, YouTube videos, TED Talks and multimedia courseware are increasingly used to expose learners to natural pronunciation, intonation and discourse. This review follows PRISMA 2020 reporting principles to synthesize peer-reviewed studies from 2020–2026 on AV materials and English speaking outcomes. Fourteen documents were assessed against pre-specified criteria, and four studies (Indonesia, Bangladesh; n = 6 to 50) were included. Every study favored AV-supported instruction for fluency, pronunciation, confidence, interaction and overall oral performance, but the evidence is weak: two studies reported scores, none had a comparison group, baseline or follow-up, and the remaining evidence is self-reported. No study measured grammatical accuracy or tested a moderator. The findings are consistent with Dual Coding Theory, Mayer's multimedia learning theory, and the Input and Output Hypotheses, yet cannot distinguish among them. The review identifies methodological gaps and recommends controlled studies that vary one component at a time (task, audio-only versus AV, captioning) with reliable speaking ratings and delayed post-tests.

***Keywords*— audiovisual materials, English speaking skills, oral proficiency, EFL/ESL learners, multimedia learning, systematic review, PRISMA**

I. INTRODUCTION

1.1 Background and Context

Speaking is widely regarded as the most demanding skill for English as a Foreign Language (EFL) and Second Language (ESL) learners, because oral production requires real-time coordination of vocabulary, grammar, pronunciation and pragmatic competence under social pressure. Limited exposure to natural spoken English, low confidence and high anxiety contribute to persistently weak speaking outcomes (Kathirvel & Hashim, 2020). AV materials are positioned as a technology-based strategy because learners can see and hear pronunciation, stress and intonation simultaneously and then imitate them, which may reduce anxiety and build confidence. Practical constraints recur in the literature: difficulty sourcing suitable materials, insufficient teacher training, and unequal access to equipment and connectivity (Kathirvel & Hashim, 2020).

Empirical work since 2020 spans many contexts. In a quasi-experiment with 80 Jordanian university students, a YouTube-based group outperformed a traditional-instruction group, with pronunciation and fluency/coherence improving most and grammatical accuracy least (Saed et al., 2021). Similar designs reported gains from authentic video among Iraqi

undergraduates (Saber, 2025) and from animation video among Indonesian tenth-graders (Fitri, 2022; η^2 = .179). Task-based action research combined AV media with communicative tasks and raised class speaking means from 45.4 to 75.8 (Pinto & Lomi, 2026). Qualitative and mixed-methods studies add teacher perspectives in Namibian primary classrooms (Shikongo, 2025), multimodal gains among 214 Iraqi students (Al-jumaily & Alazzawi, 2025), and combined listening-and-speaking gains in Bangladesh (Rahman & Jamila, 2024). These sources point in a common direction but differ widely in design and outcome measure.

1.2 Theoretical Grounding

Five frameworks guide the review. **Dual Coding Theory** (Paivio, 1971, 1986; Clark & Paivio, 1991) holds that verbal and imagery information are processed in separate but linked systems, so a voice paired with visible articulation and context offers two routes for encoding pronunciation. **Mayer's Cognitive Theory of Multimedia Learning** (2020) adds that limited-capacity channels benefit from coherent, well-segmented input with low extraneous load, so clip length and pacing matter. **Krashen's Input Hypothesis** (1985) treats comprehensible input as the driver of acquisition, with visual context making input understandable, although speaking is seen as an eventual output. **Swain's Comprehensible Output Hypothesis** (1985) argues that learners also need to produce language to notice gaps, which supports using AV material as a platform for production tasks. **Vygotsky's Sociocultural Theory** (1978) frames AV materials as cultural tools whose value depends on teacher-mediated, socially structured activity. The first two explain cognitive mechanisms, the next two offer partly competing accounts of exposure versus production, and the last situates both in social interaction.

1.3 Problem, Purpose and Research Questions

Existing evidence is dispersed, drawn from varied ages and contexts, and operationalizes "speaking improvement" in incompatible ways: IELTS band descriptors, effect sizes, action-research score bands, teacher perception and composite communication scores. Some studies bundle speaking with listening. A transparent, criteria-driven review is needed to determine which speaking dimensions are most reliably enhanced, under what conditions, and for which learners. The purpose is to identify, appraise and synthesize peer-reviewed empirical evidence from 2020–2026 on how AV materials affect English speaking among EFL/ESL learners. Four questions guide the review:

- **RQ1:** What types of AV materials and approaches (passive viewing versus active production) have been used?
- **RQ2:** Which speaking dimensions (fluency, pronunciation, accuracy, confidence, interaction) are most positively affected?
- **RQ3:** Which conditions (task integration, teacher mediation, feedback, age/proficiency, resources) moderate effectiveness?
- **RQ4:** What methodological trends and gaps characterize the research, including whether speaking can be isolated from other skills?

The review is delimited to studies published 2020–2026 that report a measured or assessed speaking-related outcome in EFL/ESL learners of any age. It aims to give curriculum designers and teachers evidence-based guidance and to consolidate a fragmented post-pandemic literature.

II. LITERATURE REVIEW AND CONCEPTUAL FRAMEWORK

2.1 Defining Audiovisual Materials

AV materials are tools that combine sound and image to make language input more concrete, authentic and memorable than text alone (Kathirvel & Hashim, 2020). For speaking, two functions are useful to distinguish. As **modeling input**, learners watch and absorb pronunciation, intonation and discourse patterns through exposure (Saed et al., 2021; Saber, 2025). As a **platform embedded in production tasks**, the material is woven into cycles of communicative activity or used as a live scaffold for imitation and role-play (Pinto & Lomi, 2026; Shikongo, 2025). The distinction maps onto the Input–Output contrast: passive-viewing designs align with an Input account, task-integrated designs with a Comprehensible Output account, and teacher-mediated use with Sociocultural Theory.

2.2 Speaking Sub-constructs

Studies rarely treat speaking as a single outcome, and results are not uniform. **Fluency** appears in nearly every study and improved markedly in YouTube-based and task-based interventions (Saed et al., 2021; Pinto & Lomi, 2026). **Pronunciation** is the construct most consistently linked to AV input, consistent with the theoretical expectation that synchronized audio-visual

Crossref DOI: https://dx.doi.org/10.22161/ijtle.5.5.2

modeling supports encoding of segmental and prosodic features. **Grammatical and lexical accuracy** appears less responsive: gains were present but smaller than for fluency and pronunciation in Saed et al. (2021), a pattern resting on a single study. **Communicative confidence** is both an outcome and a mediator; hearing correct pronunciation in advance may indirectly build confidence and reduce shyness (Kathirvel & Hashim, 2020; Pinto & Lomi, 2026), though teacher-perceived confidence (Shikongo, 2025) is not a validated measure. **Interactional competence** is rarely isolated; Al-jumaily and Alazzawi (2025) combined verbal and nonverbal communication, making it only a provisional data point.

2.3 Instructional Approaches and Prior Reviews

Four patterns appear: video as passive modeling input replacing traditional content (Saed et al., 2021; Saber, 2025; Fitri, 2022); AV embedded in task cycles (Pinto & Lomi, 2026); teacher-mediated modeling for young learners through rhymes and role-play (Shikongo, 2025); and multimodal, captioned input with bundled outcomes (Al-jumaily & Alazzawi, 2025). "AV materials for speaking" is therefore a family of approaches whose design may moderate effects. Prior systematic work has concentrated on listening: Shaojie, Samad and Ismail (2022) reviewed 46 studies on AV input and listening comprehension and, by design, excluded speaking. Speaking-focused reviews exist but are limited to YouTube and appear to have used a single database (Maulidiah et al., 2026). The present review therefore claims a narrower contribution: speaking outcomes across AV media types, with explicit eligibility rules and a quality appraisal step.

III. METHODOLOGY

3.1 Design, Protocol and Reporting

This is a systematic review reported with reference to PRISMA 2020 (Page et al., 2021) and PRISMA-S (Rethlefsen et al., 2021). Narrative synthesis with comparison tables was chosen, because scoping showed speaking measured in fundamentally different ways (rated band descriptors, eta-squared, class means, perceptions) and designs from quasi-experiments to qualitative case studies. Meta-analysis was planned only if at least five studies reported comparable between-group effects. The protocol is to be registered on the Open Science Framework [registration details to be added]; PROSPERO was not used because it is limited to health-related outcomes.

3.2 Eligibility Criteria

Criteria were defined with the PICOS framework (Table 1) and fixed before screening.

Table 1. Eligibility criteria

Element	Inclusion	Exclusion
Population	EFL/ESL learners of any age, level or proficiency	Other languages; native speakers; teacher-only studies
Intervention	AV material used to develop speaking: video, film, YouTube, animation, TED Talks, courseware, captioned video, AV aids	Audio-only (except as comparator); AV used only for other skills
Comparator	Conventional instruction, no-video, audio-only or pre/post baseline	None required; type coded
Outcomes	At least one measured or assessed speaking outcome: fluency, pronunciation, accuracy, oral score, confidence, interaction	General attitudes only, with no speaking outcome
Design	Experimental, quasi-experimental, action research, mixed-methods, qualitative	Conceptual or narrative papers; systematic reviews (used for citation chasing)
Publication	Peer-reviewed journal articles, Jan 2020 onward, full text, English	Theses, conference abstracts, preprints, books, duplicates

Boundary rules were set in advance. Theses were excluded from synthesis because they lack uniform peer review. Narrative papers were excluded but used for citation chasing. Studies with bundled outcomes (e.g., listening plus speaking) were included as partial evidence only when a speaking-specific result was separately reported. Perception-only studies were eligible if peer-reviewed and in range, but were coded as

Crossref DOI: https://dx.doi.org/10.22161/ijtle.5.5.2

perception-based and never merged with performance data. Exclusion reasons were assigned in a fixed order: date, document type, outcome.

3.3 Sources, Search and Selection

The intended strategy searches Scopus, Web of Science, ERIC and a fourth education database, with Google Scholar as a supplement and forward and backward citation chasing. Three concept blocks are combined with AND: AV media (audiovisual, video, YouTube, film, animation, multimedia, TED talk, subtitles, captions), speaking (speaking, oral proficiency, fluency, pronunciation, spoken English) and learners (EFL, ESL, English language learners), with filters for 2020–2026, journal articles and English. Full strings and a pilot against known studies will be reported in an appendix. In the present analysis, however, the evidence base is a **supplied set of 14 documents** and not the output of a database search; no PRISMA flow diagram can therefore be reported. Screening was done by one reviewer, with dual screening and Cohen's kappa planned for the full search.

3.4 Extraction, Appraisal and Synthesis

A structured form captured participants, design, AV type and function (modeling versus task-embedded), teacher mediation, feedback, instruments, evidence type (performance versus perception), results and implementation constraints. Quality was appraised with the Mixed Methods Appraisal Tool 2018 (Hong et al., 2018), reporting criterion-level ratings and not an overall score; no study was excluded on quality. Synthesis followed SWiM guidance (Campbell et al., 2020), grouping by speaking sub-construct and AV function, with vote counting based on direction of effect only. Where possible, eta-squared from two-group comparisons would be converted to Hedges' g. Sensitivity analyses excluded (a) partial or bundled studies, (b) studies with two or more "No" appraisal ratings, and (c) perception-based evidence.

IV. RESULTS

4.1 Study Selection and Characteristics

Fourteen full-text documents were assessed; four were included and ten excluded (Table 2). Seven were excluded for predating 2020, two for lacking primary data, and one for reporting listening only.

Table 2. Full-text eligibility decisions (n = 14)

Document	Decision	Primary reason
Irmawati, Ihsan, & Rasmi (2020)	Included	Speaking scores (comprehension, fluency) and observation
Naz (2023)	Included (partial, borderline)	Perceived interaction and participation; no speaking-specific measure
Rahman & Jamila (2024)	Included (partial)	Listening and speaking assessed together; speaking reported separately
Dewayanti, Artini, & Kusuma (2026)	Included	Perceived speaking, pronunciation and confidence gains from films
Kathirvel & Hashim (2020); Maulidiah et al. (2026)	Excluded	No primary data (narrative paper; systematic review)
Syah & Fediyanto (2026)	Excluded	Listening outcome only
Mizab (2015); Bahrani & Sim (2012); Bajrami & Ismaili (2016); Macwan (2015); Kim (2015); Sabiq (2017); Polat & Erişti (2019)	Excluded	Published before 2020 (several also thesis, conceptual, or non-speaking)

Of the included studies, one was published in each of 2020, 2023, 2024 and 2026. Two were from Indonesia and two from Bangladesh; two involved secondary students and two university students; none involved primary learners. Samples ranged from 6 to 50. Table 3 summarizes their features.

Crossref DOI: https://dx.doi.org/10.22161/ijtle.5.5.2

Table 3. Characteristics of included studies (n = 4)

Feature	Irmawati et al. (2020)	Naz (2023)	Rahman & Jamila (2024)	Dewayanti et al. (2026)
Context	Indonesia; Grade VIII	Bangladesh; first-year undergraduates	Bangladesh; first-year, pre-intermediate	Indonesia; Grade 10
Sample	28 (one class)	50 stated; 9 in focus group	46 (14–36 per test); 24 surveyed	6
Design	Action research, 2 cycles, no control	Survey plus focus group	Single group, 10 tests over ~3 months	Qualitative interviews
AV material	YouTube dialogues, slides	Slides, short films, clips, Mentimeter	TED Talks, BBC Learning English clips	Feature films (self-selected)
Evidence	Performance + perception	Perception	Performance + perception	Perception

4.2 Types of AV Material and Approach (RQ1)

Four different AV types appeared, one per study. Three studies combined AV material with production tasks (pair role-play, student presentations, video presentations) and teacher mediation, so effects cannot be attributed to AV alone. Only Dewayanti et al. (2026) described largely self-directed viewing, and its evidence is perception-based.

4.3 Effects by Speaking Sub-construct (RQ2)

Two studies reported scores, but scales and tasks differ and results were not pooled. Irmawati et al. (2020) reported class means rising from 57.78 to 72.67 between cycles (fluency 57.22 to 70.00); the average of the two Cycle II aspect scores is 71.67, so the gain may be 13.89, not 14.89. Rahman and Jamila (2024) reported mean test scores rising from 13.57 to 15.16 out of 20 across ten tests (+1.59 points), with dips at Tests 3, 6 and 7. Different numbers of students sat each test, tasks became harder over time, scoring criteria were not described, and neither study reported inferential statistics, so no effect sizes could be computed.

Table 4. Summary of evidence by sub-construct

Sub-construct	Studies	Evidence	Strength
Fluency	2	Irmawati: 57.22 to 70.00; Dewayanti: imitation of actors	Very limited
Pronunciation	2	Perception only (Irmawati interviews; Dewayanti)	Very limited
Accuracy	0	None	No evidence
Confidence	4	Self-reported in all four; no validated instrument	Limited, favorable
Interaction	2	Naz (about 84% disagreed multimedia is unrelated to interaction); Irmawati observation	Very limited
Overall oral performance	2	Score increases in Irmawati and Rahman & Jamila; no comparison group	Very limited

All reported findings favored AV-supported instruction, and no study reported a null or negative result. With four studies, two perception-only and none with a comparison group, this pattern cannot be tested for publication bias and should not be read as strong evidence of effectiveness. Learners also reported difficulty with fast dialogue, accents, idioms and slang, coping through subtitles and replaying scenes (Dewayanti et al., 2026), and initial unfamiliarity with AV-supported testing (Rahman & Jamila, 2024).

Crossref DOI: https://dx.doi.org/10.22161/ijtle.5.5.2

4.4 Moderating Factors (RQ3)

No moderator was examined. Task design, age, proficiency, exposure, feedback, teacher mediation and captioning were described but never compared within a study. Duration ranged from about four meetings to 24 sessions over three months. Between-study differences in context, design and outcome cannot be treated as moderator effects.

4.5 Methodological Patterns (RQ4)

Speaking was scored in 2 of 4 studies, and perception data appeared in all four. None reported rater reliability, a scoring rubric, a control group, a baseline, a follow-up or a validated confidence instrument; only Irmawati isolated speaking cleanly, and only one questionnaire reported reliability (α = .691). Reporting inconsistencies were found in every study: Irmawati's Cycle II mean does not match its aspect scores; Naz's stated N of 50 conflicts with figure percentages consistent with about 39 responses; Rahman and Jamila describe "significant" gains using only descriptive statistics; and Dewayanti's tables attribute the speaking theme inconsistently.

4.6 Quality Appraisal and Sensitivity Analysis

Provisional MMAT ratings by one reviewer are shown in Table 5 (Y = yes, N = no, CT = can't tell). No study met all five criteria, and two had two "No" ratings.

Table 5. Provisional MMAT (2018) ratings (single rater)

Study (category)	C1	C2	C3	C4	C5	Note
Irmawati (mixed)	CT	CT	CT	N	N	No baseline or comparison; concurrent changes
Naz (mixed)	Y	CT	CT	CT	N	N inconsistent with figures; AV not an intervention
Rahman & Jamila (descriptive)	CT	CT	CT	N	N	Different examinees per test; no inferential analysis
Dewayanti (qualitative)	Y	Y	CT	CT	Y	Six participants; few quotations

Excluding partial studies (Naz; Rahman & Jamila) left Irmawati and Dewayanti, which still favored AV on fluency, pronunciation and confidence, but interaction lost its only dedicated evidence. Excluding the two studies with two "No" ratings left only perception-based evidence. Excluding perception-based evidence left two studies with score increases but no comparison group. The direction never changed, but only because all four studies favored AV.

4.7 Summary Against the Research Questions

- **RQ1:** Four AV types, one each; three paired with tasks and teacher mediation.
- **RQ2:** All findings favored AV for fluency, pronunciation, confidence, interaction and overall oral performance, but evidence is limited and accuracy was not measured.
- **RQ3:** No moderator was tested.
- **RQ4:** No control group, baseline, follow-up, rater reliability or validated instrument; only two of 14 documents combined a speaking outcome, an in-range date and a performance score.

V. DISCUSSION

5.1 Summary and What the Evidence Supports

Four studies favored AV-supported speaking instruction across fluency, pronunciation, confidence and interaction. The evidence is consistent with the possibility that AV-supported instruction is associated with better performance and more positive attitudes. It does not show that AV material causes gains, that it outperforms other instruction, how large any benefit is, or which design features matter. Four studies are a small base, and uniformly positive results may reflect over-representation of favorable work.

5.2 Theoretical Interpretation

Input versus output. The expectation that task-embedded AV should outperform passive viewing could not be tested, because no study compared the two. The two score-reporting studies used task-embedded designs and both improved, which is compatible with an output account but equally with repeated practice of the same test, teacher feedback and growing familiarity. Support from other documents is conceptual (Bajrami & Ismaili, 2016; Macwan, 2015) or secondhand (Maulidiah et al., 2026, whose seven-document base includes three reviews and which excluded studies partly for lack of effectiveness).

Crossref DOI: https://dx.doi.org/10.22161/ijtle.5.5.2

Dual Coding and CTML. Synchronized sound and image plausibly give two channels for encoding pronunciation. The data are consistent with this but do not test it: the two sources on pronunciation rely on learner reports, no study compared audio-visual with audio-only input, and the claim that AV media were "almost twice" as effective as single-channel media in Irmawati et al. has no comparison condition behind it. Imitation of actors resembles shadowing, so repetition and motivation are a rival explanation. CTML's load concerns are echoed by difficulty with fast speech and by Irmawati's limit of three plays, though neither was tested. A single student's report cannot "prove" a cognitive theory, as Naz (2023) suggests.

Proficiency and mediation. Older documents hint that input difficulty matters: low-level learners did not improve with news but did with cartoons and films (Bahrani & Sim, 2012), and listening gains were larger at higher levels (Kim, 2015; Polat & Erişti, 2019). These concern other skills and only motivate hypotheses for speaking. Teacher mediation described in Irmawati et al. and Rahman and Jamila fits a sociocultural framing, but again formed part of a package. Overall, all five theories predict the same pattern, so none is favored; distinguishing them requires designs that vary one component at a time.

5.3 Comparison with Earlier Evidence

Listening reviews (Shaojie et al., 2022) found that subtitles, video type and sound–image relation shape difficulty; among included speaking studies only subtitle use appeared, as a coping strategy. Supplied listening studies had stronger designs (randomized groups, validated anxiety scales) than any included speaking study, supporting the premise that the speaking base is less developed. Maulidiah et al. (2026) concluded that YouTube is effective, especially with structured tasks, but their single-source search and exclusion of studies for "lack of effectiveness" may favor positive findings. Perception studies are consistently favorable across years: 77.5% of students saw AV aids as important for speaking (Mizab, 2015) and 83.4% rated AV materials effective for listening and speaking (Rahman & Jamila, 2024), alongside a persistent gap between endorsement and classroom provision. Attitudes, however, are not evidence of learning. Action-research designs also tend to show cycle-to-cycle gains whatever the skill (Sabiq, 2017), limiting what Irmawati's gain can say about AV specifically.

5.4 Pedagogical Implications

These suggestions are reasonable practice, not effect-size findings.

- **Pair AV input with production tasks,** such as role-play from a dialogue video or a video presentation.
- **Plan teacher mediation:** limit plays, simplify instructions, use gesture instead of translation, and grade tasks from easy to complex.
- **Select materials against explicit criteria:** learner level, speech rate, specialized vocabulary, accent variety, and subtitle availability.
- **Keep the audio in view,** using shadowing or imitation of short segments rather than attending only to visuals (Kathirvel & Hashim, 2020).
- **Support offline access** through pre-downloaded clips, and **align assessment** so that speaking is regularly assessed and AV-based practice keeps its priority.

5.5 Gaps in the Literature

- **Longitudinal evidence:** no delayed post-test; the longest study ran about three months.
- **Instruments:** no validated confidence or anxiety scale, no reported rater reliability or rubric.
- **Populations:** no primary-school learners; only Indonesia and Bangladesh represented.
- **Accuracy and higher-order skills:** accuracy, interactional competence and discourse organization were unmeasured; the claim that grammar benefits less than fluency rests only on a secondhand summary of Saed et al. (2021).
- **Controlled comparisons:** no AV versus audio-only, passive versus task-embedded, or captioned versus uncaptioned contrasts.

5.6 Limitations

The 14 documents were supplied rather than retrieved through a database search, so selection was not systematic and the PRISMA flow cannot be reported. Only English full texts were considered, theses were excluded by rule (removing relevant data such as Shikongo, 2025), and seven documents were excluded for predating 2020. All four included studies, and all documents in Maulidiah et al. (2026), favored AV, so publication bias cannot be excluded. Samples were small and heterogeneous, pooling was impossible, and findings are descriptive. Screening, extraction and MMAT appraisal were done by one reviewer, and the

Crossref DOI: https://dx.doi.org/10.22161/ijtle.5.5.2

borderline inclusion of Naz (2023) could be reversed (leaving three studies without changing the direction of findings). Some included articles contain internal inconsistencies, and the peer-review status of some journals remains to be confirmed.

5.7 Conclusion

The literature tells a consistent story about attitudes and a cautiously positive story about speaking performance, but it is thin on evidence that AV material, as distinct from good teaching, is responsible. The most useful next step is a small set of controlled studies that deliver the same AV material with and without a production task, compare audio-visual with audio-only input, and use reliable ratings of fluency, pronunciation and accuracy with delayed follow-up.

REFERENCES

[1] Al-jumaily, I. H. A., & Alazzawi, I. T. J. (2025). The influence of multimodal visual methodologies on EFL university students' audio-visual comprehension, verbal and nonverbal communication. *Journal of Language & Education, 11*(2).†

[2] Bahrani, T., & Sim, T. S. (2012). Audiovisual news, cartoons, and films as sources of authentic language input and language proficiency enhancement. *The Turkish Online Journal of Educational Technology, 11*(4), 56–64.

[3] Bajrami, L., & Ismaili, M. (2016). The role of video materials in EFL classrooms. *Procedia – Social and Behavioral Sciences, 232*, 502–506. https://doi.org/10.1016/j.sbspro.2016.10.068

[4] Campbell, M., McKenzie, J. E., Sowden, A., et al. (2020). Synthesis without meta-analysis (SWiM) in systematic reviews: Reporting guideline. *BMJ, 368*, l6890.

[5] Clark, J. M., & Paivio, A. (1991). Dual coding theory and education. *Educational Psychology Review, 3*(3), 149–170.

[6] Dewayanti, N. P. W., Artini, L. P., & Kusuma, I. P. I. (2026). Exploring the role of English movies in language learning: Senior high school students' perceptions. [Journal title to be confirmed], 5(8), 1322–1333.†

[7] Fitri, R. D. (2022). *The effect of using animation video as audio-visual media toward students' speaking skill at the tenth grade of SMAN Plus Riau Province* [Undergraduate thesis, State Islamic University of Sultan Syarif Kasim Riau].

[8] Hong, Q. N., Pluye, P., Fàbregues, S., et al. (2018). *Mixed Methods Appraisal Tool (MMAT), version 2018*. Canadian Intellectual Property Office.†

[9] Irmawati, Ihsan, M., & Rasmi. (2020). Utilizing audio visual aids to improve English speaking skill for the eight grade students of SMP Negeri 3 Banawa. *Datokarama English Education Journal, 1*(1), 40–50.†

[10] Kathirvel, K., & Hashim, H. (2020). The use of audio-visual materials as strategies to enhance speaking skills among ESL young learners. *Creative Education, 11*(12), 2599–2608. https://doi.org/10.4236/ce.2020.1112192

[11] Kim, H.-S. (2015). Using authentic videos to improve EFL students' listening comprehension. *International Journal of Contents, 11*(4), 15–24.

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

[13] Macwan, H. J. (2015). Using visual aids as authentic material in ESL classrooms. *Research Journal of English Language and Literature, 3*(1), 91–96.

[14] Maulidiah, F., Kurrata'Aini, R., Yanti, M., & Martiarini, E. (2026). The effectiveness of using YouTube for teaching EFL in improving student's speaking skills: A systematic literature review. *Indonesian Research Journal on Education, 6*(2), 191–198.

[15] Mayer, R. E. (2020). *Multimedia learning* (3rd ed.). Cambridge University Press.

[16] Mizab, M. (2015). *The role of audio-visual aids in enhancing EFL students' speaking skill* [Master's dissertation, Mohamed Kheider University of Biskra].

[17] Naz, S. (2023). Role of multimedia-aided EFL classrooms in promoting learners' interaction and participation in tertiary-level Bangladeshi students. *Journal of Language Teaching and Research, 14*(5), 1207–1214. https://doi.org/10.17507/jltr.1405.08

[18] Page, M. J., McKenzie, J. E., Bossuyt, P. M., et al. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *BMJ, 372*, n71.

[19] Paivio, A. (1971). *Imagery and verbal processes*. Holt, Rinehart & Winston.

[20] Paivio, A. (1986). *Mental representations: A dual coding approach*. Oxford University Press.

[21] Pinto, M. A. M., & Lomi, A. N. K. (2026). Integrating task-based activities and audio-visual media to enhance EFL students' speaking ability. *Ebony, 6*(1), 78–97. https://doi.org/10.37304/ebony.v6i1.23891

[22] Polat, M., & Erişti, B. (2019). The effects of authentic video materials on foreign language listening skill development and listening anxiety at different levels of English proficiency. *International Journal of Contemporary Educational Research, 6*(1), 135–154.

[23] Rahman, M. M., & Jamila, M. (2024). Effectiveness of audiovisual materials in developing tertiary level learners' English listening and speaking skills. *International Journal of Language Testing, 14*(2), 67–81. https://doi.org/10.22034/IJLT.2024.430050.1312

[24] Rethlefsen, M. L., Kirtley, S., Waffenschmidt, S., et al. (2021). PRISMA-S: An extension to the PRISMA statement for reporting literature searches in systematic reviews. *Systematic Reviews, 10*, 39.

[25] Saber, A. F. (2025). Examining the effectiveness of authentic video materials in developing speaking skills among undergraduate EFL Iraqi learners. *Journal of Tikrit University for Humanities, 32*(11–2), 22–37.

[26] Saed, H. A., Haider, A. S., Al-Salman, S., & Hussein, R. F. (2021). The use of YouTube in developing the speaking

skills of Jordanian EFL university students. *Heliyon, 7*(7), e07543.

[27] Sabiq, A. H. (2017). Improving students' motivation and reading skills in ELT through audio visual media. *Diksi, 25*(1).† (Cited in text as "As Sabiq".)

[28] Shaojie, T., Samad, A. A., & Ismail, L. (2022). Systematic literature review on audio-visual multimodal input in listening comprehension. *Frontiers in Psychology, 13*, 980133.

[29] Shikongo, M. E. P. (2025). *Exploring teachers' use of audio-visual aids in teaching English-speaking skills to junior primary learners* [Master's thesis, University of Namibia].†

[30] Swain, M. (1985). Communicative competence: Some roles of comprehensible input and comprehensible output in its development. In S. M. Gass & C. G. Madden (Eds.), *Input in second language acquisition* (pp. 235–253). Newbury House.

[31] Syah, D. A., & Fediyanto, N. (2026). Animated videos in English listening skill development. *Academia Open, 11*(1).

[32] Vygotsky, L. S. (1978). *Mind in society*. Harvard University Press.



International Journal of Teaching, Learning and Education (IJTLE)

ISSN: 2583-4371

Vol-5, Issue-5, September-October 2026

Journal Home Page: https://ijtle.com/

Journal DOI: 10.22161/ijtle

Transforming Higher Education through Digital Technologies: A Critical Analysis

Dr. Sheeba Parveen

Assistant Professor, Department of English, R.B.D. Mahila Mahavidyalaya, Bijnor, UP, India

Received: 18 Sep 2026, Received in revised form: 11 Oct 2026, Accepted: 15 Oct 2026, Available online: 20 Oct 2026
©2026 The Author(s). Published by IJTLE. This is an open-access article under the CC BY license (https://creativecommons.org/licenses/by/4.0/).

***Abstract*—** *"Digital technologies are transforming the way knowledge is produced, accessed, and shared in higher education."*

— Neil Selwyn, Education and Technology: Key Issues and Debates

Digital transformation has revolutionised the world of higher education. In contemporary time period digital technology such as artificial intelligence, online learning platform, learning management systems and data analysis have changed the teaching learning environment in colleges and universities throughout the world. Digital technology is used for administrative activities as well. This article examines the challenges and opportunities of digital transformation in the world of higher education. Drawing upon other researches, conducted by Neil Selwyn and George Siemens, this study tries to identify different ways in which digital technology increases access to education, facilitates learning that is customised and flexible, facilities Global learning and improve efficiency within educational institutions. This paper also examines some other issues such as data privacy and insecurity concerns, digital gap, inadequate digital literacy, infrastructure constraints and other such issues.

Keywords*— *Digital Transformation, Higher Education, Educational Technology, Online Learning, Digital Divide.

"The integration of technology in education has fundamentally reshaped teaching and learning processes."

— UNESCO Report

Digital transformation is a concept that describes the use of digital technologies on every facet of education, which is a fundamental way of altering the functions and provision of learning within institutions. Digital transformation in higher education encompasses online learning systems, artificial intelligence, data analytics, virtual learning classrooms, and digital administration systems.

The sudden increase in the number of digital technologies, in particular, following the COVID-19 pandemic has increased the implementation of digital tools in the world universities. Although this change opens many opportunities that can be harnessed to enhance learning opportunities, access, and efficiency, several limitations have been created due to the change, including the issue of digital inequality, infrastructure, and pedagogy.

Alongside multiple advantages, there are also a number of challenges and restrictions of digital transformation. The most important problem is probably the digital divide implying the inequalities in access to the digital technologies amongst students. Not all the students, especially those in developing nations or in the rural regions will have reliable internet connections and digital devices. This disparity may be a hindrance to learning and increase the divide between the fortunate and unfortunate students. The second limitation is that not many universities have sufficient technological infrastructure. Digital learning solutions

are easily implemented in institutions that have limited internet bandwidth, due to the outdated technological systems, and lack of funds. The digital change can't be completely achieved without the appropriate infrastructure.

The problem of the faculty resistance and the discrepancy in skills of the staff are also a barrier to the effective use of digital technologies in education. Other teachers might be reluctant to embrace the digital use because of the lack of technological literacy or intimidation with change. As such, the institutions will have to offer training and professional development courses to assist teachers to acquire digital skills and accommodate the novel teaching practices.

Another aspect that matters in the context of digital transformation is the security of the Internet and privacy of data. Schools and colleges gather and keep extensive records of students in the digital format and are a susceptible target of cyber attacks and data breaches. The security and privacy of academic data is the key to ensuring the trust in the digital learning systems. Another challenge is to ensure the quality and integrity of academic work in online education. The plagiarism problem, cheating in online tests and lower involvement of students can threaten the reputation of distance learning. To ensure that academic standards are not compromised in the virtual learning setting, universities need to build viable monitoring and evaluation plans. The last one, digital transformation usually involves a huge financial expenditure. Universities will be forced to spend on technological infrastructure, software systems, cyber security infrastructure, and faculty development. Such investments may not be economically feasible to many institutions particularly in the developing countries.

To ensure that they reap maximum gains of digital transformation, institutions of higher learning need to be strategic and inclusive. Among these strategies, one of the primary ones is investing into the durable digital infrastructure, which is the high speed of internet connection, progressive learning platforms, and stable technological tools. These investments are necessary to promote digital learning and teaching processes.

Continuous facility training programs are another strategy that is important. The teachers should be efficient in using digital technology for making themselves empowered. They should have pedagogical knowledge of how to use technology in teaching an learning process. Professional development programmes should be introduced to train teachers or educators in digital teaching learning skills. It has become very essential that institutions should introduce inclusive digital policies and put efforts in the direction of decreasing the digital divide. The ideal world in which every student get easy access to digital education, can become a reality, when there is affordable internet connections and technical assistance available. In the process of digital transformation, cyber security enhancement also plays a very significant role. There is need of some initiatives such as : cyber security training programs and establishing secure digital platform to promote integrity of digital systems.

Lastly, balanced and effective learning can be developed by adopting blended learning models that integrate both the traditional classroom-based learning with the online learning. Blended learning enables the institutions to combine the benefits of physical and online learning systems.

The future of the higher education will be tied more to the emerging technologies that can revolutionize the learning environment. Such technologies as artificial intelligence, virtual reality, augmented reality, and block chain will have a huge impact on the future of education. Such inventions have the potential to generate second-degree and online learning that improves student participation and comprehension. AI can be used to facilitate personalized learning and intelligent tutoring systems, and virtual and augmented reality could be used to replicate real-world experiences to facilitate practice learning. Another area of block chain technology application is the use of secure and transparent academic credentials storage system and the verification of qualifications. Moreover, the possibilities of learning in the digital realm can be extended with the creation of multiple ways based learning environments, which will allow building virtual campuses where students and teachers can communicate in immersive digital environment.

The concept of digital transformation is significantly changing the face of higher education in the world. It provides immense chances to increase the level of accessibility, better modes of teaching, and train students in accordance with the requirements of the digital economy. Simultaneously, it has a number of issues associated with digital disparity, insufficient infrastructure, cyber security threats, and faculty preparedness.

To ensure success in digital transformation, universities should have a holistic strategy to

Crossref DOI: https://dx.doi.org/10.22161/ijtle.5.5.3

incorporate technological innovation with good pedagogical practices and accommodative institutional policies. If the colleges and universities get success in overcoming the challenges and invest in digital technology, infrastructure and skill development programmes, it will result in developing an educational system that will be flexible, innovative and resilient.

REFERENCES

[1] Adeyemi, C., Roberts, M., & Hernandez, S. (2025). Digital Transformation in Higher Education: Opportunities and Challenges.
[2] Yadav, R. U. (2025). Digital Transformation in Higher Education: Challenges, Opportunities and Future Directions.
[3] Gkrimpizi, T., & Peristeras, V. (2023). Digital Transformation in the Higher Education Sector: A Systematic Literature Review.
[4] UNESCO (2021). Digital Transformation of Education.
[5] Bates, A. W. (2020). Teaching in a Digital Age.
[6] Selwyn, Neil. Education and Technology: Key Issues and Debates. 2nd ed., Bloomsbury Academic, 2016.

Crossref DOI: https://dx.doi.org/10.22161/ijtle.5.5.3

International Journal of Teaching, Learning and Education (IJTLE)
ISSN: 2583-4371
Vol-5, Issue-5, September-October 2026
Journal Home Page: https://ijtle.com/
Journal DOI: 10.22161/ijtle

Contribution of School Management Strategies to Student's Academic Performance in Primary School in Misungwi District, Tanzania

Christopher John Legonda

Prime Minister's Office - Regional Administration and Local Government, Education Division, Pre and Primary Education, Ilemela Municipal Council, P. O. Box 735, Mwanza, Tanzania.
*Corresponding author Email: legondac@yahoo.com

Received: 15 Sep 2026, Received in revised form: 10 Oct 2026, Accepted: 16 Oct 2026, Available online: 20 Oct 2026
©2026 The Author(s). Published by IJTLE. This is an open-access article under the CC BY license (https://creativecommons.org/licenses/by/4.0/).

Abstract

The major role of leadership aspects in influencing the academic outcomes of the students is based upon the administration and management of the school. Management of education particularly in primary schools takes into deliberation the application of management policies in designing, developing and producing resources towards achievement of students' academic performance. There is a growing body of information from various scholars which have made an attempt to scrutinize the effect of school management and the academic performance of students which support this study. This study ascertains the Effective Supervision of School Management Strategies in Contributing to Student's Academic Performance in Primary School. The primary data were collected using questionnaire tool from 256 respondents. The sample was chosen from each selected Primary School by simple random sampling technique. The data were analyzed using Statistical Package for Social Science (SPSS ver. 20, IBM, USA). The findings revealed that school management take necessary actions and delegation of power as the school administrative strategies which contribute to students' academic performance. The study recommends that, planned school management strategies and their implementation should be promptly communicated to all consulting groups and stakeholders such as students, teachers, parents and other employees from all levels.

Keywords*— *Administration, Management, Academic performance, School, Head of School

I. INTRODUCTION

School management strategies play's a crucial role in the education sector development worldwide, because influence students' academic performance (Khuzwayo, 2007 & Amadi & Edu, 2020). Strong management strategies significantly improve educational and student academic performance (Kemenanabo, 2019). Strength of school's management strategies impact discipline and order among teachers which improve their performance since poor performance is considered unacceptable and may be viewed as misbehaviour (Islam and Tasnim, 2021).

In a modern world, the educational sector has experienced a drastic increase in school management responsibilities with the rise of global economic pressure (World Bank, 2018 & Ezeh & Ogara 2020). School management play generous roles in improving the quality of education around the world (UNESCO, 2024). In Europe such as Germany the development and implementation of management control systems

gained recognition after achieving success in the US in the early 20th century and in Europe in the second half of the 20th century (Dimov and Iliev, 2010).

The initiatives to enhance the quality and efficacy of education in Sub-Saharan Africa is the effort to improve school management. Example South Africa has provided leadership training for school administrators, which focus on improving the quality of instruction in line with Senegal's Improvement Plans (SIP) which was established in 2014. Also, Kenyan government has shown a commitment to improve the quality of education. Among of the strategies used is administrators' improvement and establishment of a department responsible for quality control, and enhancements in the managerial aspects of the education system in the country (Republic of Kenya, 2014 & Turyatunga *et al.,* 2023). In Tanzania school management strategies includes performing different functions in order to manage students' academic performance; these functions include control students' discipline, involving parents in students' learning process and improving teaching and learning environment (Mabula and Mkulu, 2022). Academic performance serves as a benchmark for effectiveness of school management strategies, which is crucial in motivating teachers to be effective in their instruction, ultimately impacting students' achievements (Omari and Kiwonde, 2022)

Kalagbor, (2016) revealed that students in private schools are more positively influenced in their academic performances by most of the factors namely: Availability of school facilities and the extent they are put in to use, friendly relationship between the teachers and students, prompt complaint to school teachers' welfare needs including payment of salaries, students' high parental status, principals' and teachers' good working relationships, teachers' supervision is prompt and effective, required professional and qualified teachers, effective utilization of lesson periods, teachers not engaging in deviant conducts and number of students in each classroom ranges from 1-50 than students in the public schools, although, students were better influenced by the quality of teaching manpower than their private schools counterpart.

Islam and Tasnim, (2021) shows that regular attendance, regular study, hard work, dedication and self-Confidence and support by family members and others have a high impact on the students' academic performance. The main areas that have been taken into account include, factors influencing the academic performance of the students, academic performance and large number of students in class, parental and associated factors relating to academic achievement, contribution of school factors towards the academic performance of students, influence of poverty on academic achievement of students and other causes of low academic achievement (Eze, Sunday & Inegbedion, 2015).

Idd and Kiwonde (2022) revealed that the government should improve the working conditions for teachers in public schools through the following: improve working conditions by building more teachers' houses with availability of services such as electricity and water, increasing teachers' salaries and annual increments, timely teachers' promotion, building of laboratories with equipment as well as provision of mobile laboratories to enhance effective teaching and learning process, improving classrooms conditions and teaching facilities to facilitate easy teaching-learning process and laying down education policies that enhance teachers' motivation.

Tanzania education policy has been implemented in ensuring that school management perform their duties as planned. The focus of education policy has been to ensure school management strategies are in a good control in solving school students related problems and hereby improving academic performance (URT, 2024). Despite the potential benefit of school management strategies; there is limited empirical evidence on the contribution of school managerial strategies towards students' academic performance in Misungwi District. The lack of research insight limits and questions the assessment of management strategies contribution towards students' academic performance in primary school in Misungwi District. The study therefore, aimed at assessing the contribution of management strategies towards students' academic performance in primary schools in Misungwi District for the purpose of drawing practical lesson of good students' academic performance since the school management strategies spells out good students' academic performance of which primary schools in Misungwi District is non-exceptional.

Theory Guiding the Study

Transformational leadership enhances the motivation, morale and performance of followers through a variety of mechanisms. The transformational leadership theory, was developed by James V. Downton in 1973, serves as the guiding framework for this study. Transformational leadership theories emphasize the

Crossref DOI: https://dx.doi.org/10.22161/ijtle.5.5.4

relationships formed between leaders and their followers. This hypothesis is often compared to charismatic leadership theories, which suggest that leaders who possess specific traits, such as confidence, extroversion, and positive values, are better at motivating followers in this case, teachers.

Transformational leadership theory assumes a variety of roles, including coaching or mentoring, delegation, effective communication, follower development, and control. Promote collaboration and strengthen group cohesion (Jacob & Thilagaraj, 2022). In the context of this study, the theory tactics associated with transformational leadership are expecting to have a significant impact on the study subject matter (Jacob & Thilagaraj, 2022). The transformational leadership theory is selected to guide the study since it centers on how a head of school places emphasis on elements that foster the potential growth of teachers, motivating them as integral members of a team working in achieving academic goals. Head schools assess the potential of their subordinates to help identify and realize the school's vision and mission. The transformational leadership theory principles emphasize that the leadership process is geared toward ensuring that, the subordinates uphold the ethical values and responsibilities.

The transformational leadership theory has been criticized by some scholars by for favoring elite subordinates and possibly undermining democratic principles. Also, may disqualify individuals who do not seem to possess the relevant personality traits, as leaders in organizations may have significant influence over the organization's destiny by dictating how employees should behave and function. Although the theory weakness highlighted but still have strong application towards this study subject matter.

The theory is relevant to this study since it addresses the effectiveness of managerial strategies that can trigger control techniques practiced by heads of schools to enhance students' academic performance. The transformational leadership style includes motivating teachers, stimulating intellectual growth, providing in-service training, ensuring the recruitment of qualified teachers, supervising teachers, and regulating teaching and learning processes within a given school. This theory directly related to the study by highlighting the connection between the management factors used by the head of schools in private primary schools in Misungwi District, which ultimately leads to the improvement of students' academic performance. Subsequently, the use of transformational leadership theory reveals head of school as leaders who exert significant influence impacting their teachers and fostering higher morale and commitment, all in the pursuit of enhancing students' academic performance.

II. MATERIAL AND METHODS

The study employed simple random sampling whereby each person has an equal chance of being selected and the required sample size of respondents was proportionally selected from each selected Primary School. The sample size involved 256 respondents. The sample size was calculated based on the formula described by Berenson et al. (2002). Data were collected from Primary School: mainly Teachers, Students, and Head of Schools using interviews questionnaire. Both descriptive and quantitative analysis were employed. Statistical Package for Social Science (SPSS ver. 20, IBM, USA) was used to statistics findings.

III. RESULTS AND DISCUSSIONS

School Management Strategies in Ensuring Students' Academic Performance

The result from Table 13 shows that F-test is statistically significant implying that it is implausible to find out that the factors built-in in the model do not clarify the variations obtained in students' academic performance. The coefficients of decision making, motivation and parent's involvement are significant at 5% this implies that decision making, motivation and parents involvement increases students' academic performance. However, government policy and infrastructure are also positive but not statistically significant, signifying a straight relationship between these variables and students' academic performance. Training shows negative coefficients and statistically not significant to indicate inverse relationship to the students' academic performance.

Crossref DOI: https://dx.doi.org/10.22161/ijtle.5.5.4

Table 1: Regression results: School Management Strategies

Variable name	Coefficients	Standard Error	T-Stat	P-value
Motivation	0.880	0.224	4.352	0.000
Government policy	0.024	0.086	1.607	0.109
Training	-0.034	-0.041	-0.771	0.441
Infrastructure	0.011	0.010	0.194	0.846
Decision making	0.021	0.165	3.176	0.002
Parents involvement	0.015	0.105	2.005	0.046
Constant	0.150		3.711	0.000

F-value 6.893

During the interview sessions, qualitative findings collected from respondents seem to support this view.

One of the respondents from school A had the following views to say:

> "Motivation is a critical aspect in improving performance since is one of the mechanisms that enable teachers and students to raise interest and morale for work and learning. This strategy leads them to make efforts to improve performance in which can be motivated by the results and the rewards provided to them" (Interviewee, 2026).

Another respondent also added,

> "I realized that motivation is an energetic mechanism which can be applied by school management in controlling teaching and learning activities. When you motivate your teachers fairly, you will witness how they work diligently without the need for coercion. Similarly, when students are motivated, they become more competitive in their classes, striving for higher performance to attain the rewards. I have personally implemented this approach, and the majority of both teachers and students work very diligently" (Interviewee, 2026).

The study findings implies that effective implementation of decisions, governments' policies, effectively planning and student's motivation have significance influence to the students' academic performance due to effective supervision by the school management. Also, decision making, motivation and parent's involvement are significant which implies that decision making, motivation and parents involvement increases students' academic performance.

The study is in line with Idd and Kiwonde (2022) who revealed that motivating factors tend to influence teachers' commitment to their teaching and learning processes, therefore promoting students' academic performance. Also, Johnson (2017) revealed that teachers can increase students' motivation to learn while students may have an innate desire to learn, the external support provided by the teacher has a significant impact on students' learning. The teacher's role in motivation includes, creating an environment conducive to learning. Teacher's roles in encouraging support of students' autonomy, relevance and relatedness of the material increases motivation to learn.

Furthermore the study is supported by Kalagbor, (2016) who revealed that effective supervision of school management strategies is possible if there is availability of school facilities and the extent they are put into use, friendly relationship between the teachers and students, prompt complaint to school teachers' welfare needs, including payment of salaries, parental support, principals' and teachers' good working relationships, teachers' supervision is prompt and effective, required professional and qualified teachers, effective utilization of lesson periods,

Crossref DOI: https://dx.doi.org/10.22161/ijtle.5.5.4

teachers not engaging in deviant conducts and number of students in each classroom.

School Administrative Strategies in Contributing to Students' Academic Performance

The results from Table 2 show that F-test is statistically significant implying that it is plausible to conclude that there is strong evidence that school administrative strategies influence students' academic performance in a good aspect.

Table 2: ANOVA

Model	Sum of Squares	Df	Mean Square	F	Sig.
Regression	5.240	5	1.048	6.057	.000
Residual	43.256	250	.173		
Total	48.496	255			

Table (3) indicates 0.003 and 0.000 as significance responses from the participants which confirm that school management always take necessary actions and delegation of power is the school administrative strategies which contribute to students' academic performance. However, School management leaving decisions to be made by teachers, subordinates solve their problems in cooperation and school management does not interfere good decisions are also positive but not statistically significant, signifying a straight relationship between these variables and student academic performance.

Table 3: School administrative strategies

Model	Unstandardized Coefficients		Standardized Coefficients	T	Sig.
	B	Std. Error	Beta		
(Constant)	.856	.142		6.038	.000
Take necessary actions	-.164	.055	-.276	-2.999	.003
Decisions made by teachers	.100	.064	.147	1.557	.121
Subordinates solve their Problems in cooperation	.022	.028	.049	.769	.443
Not interfere good decisions	.019	.030	.040	.649	.517
Delegation of power	.410	.103	.241	3.962	.000

Table 4: Administrative ANOVA

Model		Sum of Squares	Df	Mean Square	F	Sig.
	Regression	12.000	6	2.000	13.645	.000
	Residual	36.496	249	.147		
	Total	48.496	255			

The study findings show that providing conducive environment, evaluating and Supervision, and assigning subjects to teachers is the administrative strategies which have strongly significance towards influencing students' academic performance. This is evidenced by the level of significance of 0.000, 0.000 and 0.010 respectively which is smaller than 0.05 (Table 8b). However, there is other variable which are positive namely: boosting morale, providing professional guidance and providing guidance but not statistically significant, suggesting a straight relationship between these variables and student academic performance.

Crossref DOI: https://dx.doi.org/10.22161/ijtle.5.5.4

Table 5: Coefficients

Model	Unstandardized Coefficients		Standardized Coefficients	T	Sig.
	B	Std. Error	Beta		
(Constant)	.838	.271		3.087	.002
Assigning subjects to teachers	.262	.101	.182	2.596	.010
Providing guidance	.041	.039	.066	1.059	.290
Providing Professional Guidance	.163	.121	.082	1.346	.179
Evaluating and Supervision	.134	.029	.282	4.703	.000
Providing conducive environment	-.383	.069	-.379	-5.517	.000
Boosting morale	.052	.030	.115	1.748	.082

Qualitative findings obtained from interviews with respondents appear to converge with the quantitative results, indicating a significant relationship of the administrative strategies implemented by school management regarding students' academic performance.

For-instance, one respondent said that,

> "Many of the strategies employed by school administrative are effective, even though their effectiveness can vary depending on the specific strategy used. For instance, when heads of schools rely heavily on supervision and controlling instead of motivation, it can demotivate teachers and subsequently lead to lower performance. Thus, situations such as cross-checking students' notes and classroom teaching, motivation is often more effective than warnings" (Interviewee, 2026).

Another participant if FGD said that:

> "Among of the administrative strategy which ensure students' academic performance is providing guidance in cross-checking students' as it enables the correction of students' mistakes and errors. However, many teachers encounter difficulties in effectively cross-checking students' notes, often citing reasons such as large class sizes and a heavy teaching workload. The strategies currently in use have not effectively addressed this issue, resulting in students carrying mistakes and errors from one class to another" (Interviewee, 2026).

The study findings revealed that school management take necessary actions and delegation of power as the school administrative strategies which contribute to students' academic performance. However, School management leaving decisions to be made by teachers, subordinates solve their problems in cooperation and school management does not interfere with good decision. In this matter school management ensures administrative strategies stand as the proper way of running up daily activities in order to increase students' academic performance. Also, the study findings show that providing conducive environment, evaluating and Supervision, and assigning subjects to teachers is the administrative strategies which have strongly significance towards influencing students' academic performance.

Furthermore, qualitative findings obtained from interviews with respondents appear to align with the quantitative results, indicating that strategies employed by school administrative are effective, even though their effectiveness can vary depending on the specific strategy used. For instance, when heads of schools rely heavily on supervision and controlling instead of motivation, it can demotivate teachers and subsequently lead to lower performance. The findings are in line with Maina, (2010) who revealed that the principal, teachers and the administrative staff members of the school are vested with the authority to

Crossref DOI: https://dx.doi.org/10.22161/ijtle.5.5.4

implement the managerial functions of planning, organizing, controlling, and directing the activities. The principal primarily has the right to make decisions. The decisions can be made either at one's own discretion or they may consult others and seek ideas and suggestions from teachers and other staff members of the school. The major role of leadership aspects in influencing the academic outcomes of the students is based upon the administration and management of the school. When there are proper rules, policies and management is put into practice in an appropriate manner, then there would be improvement in academic performance of the students.

IV. CONCLUSION AND RECOMMENDATION

The study findings revealed that, when school management apply a strategy of paradigm shift system such as take necessary actions leave decisions to be made by teachers, solving problems in cooperation and not interfering good decisions creates a room for all stakeholders to take active part on matters concerning the students' academic performance prosperity. Planned school management strategies and their implementation should be promptly communicated to all consulting groups and stakeholders such as students, teachers, parents and other employees from all levels. Government policy makers should enhance the application of participatory ways in implementation of school management strategies in order to accelerate the implementation of different strategies concerning development and improvement of students' academic performance.

REFERENCES

[1] Islam, A & Tasnim S., (2021), an Analysis of Factors Influencing Academic Performance of Undergraduate Students: A Case Study of Rabindra University, Bangladesh (RUB), *International Journal of Education*, 9 (3):127-135. DOI: https://doi.org/10.34293/education.v9i3.373.

[2] Johnson D., (2017). The Role of Teachers in Motivating Students to Learn, *BU Journal of Graduate Studies in Education*, 9(1): 46-49

[3] Kalagbor L.D, (2016). An Analysis of Factors Influencing Students' Academic Performance in Public and Private Primary Schools in Rivers State-Nigeria, *Journal of Education and Practice,* 7(28): 96-101.

[4] Omari, R. & Kiwonde, F. M. (2022). Motivating Factors Influencing Public Primary School Teachers' Performance in Tanzania: A Case of Mbogwe District. *Journal of Research Innovation and Implications in Education*, 6(2): 352 – 359.

[5] Kemenanabo, L. D. U. (2019). Influence of selected management factors on academic performance of secondary school students in Rivers State. International Journal of Innovative Education Research, 7(4):39-49.

[6] Dimov, O., & Iliev, P. (2010). Controlling: A modern system of management and control, Facta Universitatis-series: Economics and Organization, 7(2), 253-262.

[7] Omari, R. & Kiwonde, F. M. (2022). Motivating Factors Influencing Public Primary School Teachers' Performance in Tanzania: A Case of Mbogwe District. Journal of Research Innovation and Implications in Education, 6(2): 352 – 359.

[8] Ezeh, C. R & Ogara, P. I. (2020). Impact of Financial Management on Effective School Administration in Enugu Education Zone, British International Journal of Education and Social Sciences, 7(6): 1-15.

[9] United Nations Educational, Scientific and Cultural Organization (UNESCO), (2014). UNESCO roadmap for implementing the global action program on education for sustainable development.

[10] World Bank (2018). Realizing the promise of education for development emerging themes of World Development Report 2018. https://thedocs.worldbank.org/en/doc

[11] Jacob D.K & Thilagaraj A., (2022), Transformational Leadership Theory – A Critical Analysis with reference to Banking Sector, Webology, 19(2): 152-159

[12] Eze, Sunday C & Inegbedion H., (2015) "Key Factors Influencing Academic Performance of International Students in UK Universities: A Preliminary Investigation." British Journal of Education, 3(5): 55-68.

[13] Amadi, M. A. & Edu, D. I. (2020). Effective control system in the management of educational institution. International Journal of Institutional Leadership, Policy and Management. 2: 246-261.

[14] Khuzwayo, S. (2007). Role Relationships of School Governing Body Chairpersons and Principals in School Governance in Selected Primary and Secondary Schools in the Kwamashu area. Unpublished MEd dissertation. Durban: University of KwaZulu-Natal.

[15] Republic of Kenya, (2014). The Policy Framework for Re-aligning education to the Constitution 2014 and Vision 2030 and beyond. Ministry of Education Science and Technology.

[16] Turyatunga M, Kaziro N, Ariyo G.K 3, Kobusingye P, Kamugisha N & Friday C., (2023). The Effect of Teacher's Motivation on Academic Performance of Students in Kole District a Case Study of Alito Sub-County, International Journal of Academic Multidisciplinary Research (IJAMR), 7(2): 64-72.



International Journal of Teaching, Learning and Education (IJTLE)
ISSN: 2583-4371
Vol-5, Issue-5, September-October 2026
Journal Home Page: https://ijtle.com/
Journal DOI: 10.22161/ijtle

AI-Customized Digital Picture Books to Enhance Vocabulary Learning in Early ELT Classrooms: A Quasi-Experimental Study

Banoth Suresh[1], Dr. Narendar. N[2]

[1]Ph.D. In English Language Education, The English and Foreign Languages University, Hyderabad India.
[2]Principal of Vidyarthi College of Education, Mavala, Adilabad, Telangana, India

Received: 19 Sep 2026, Received in revised form: 15 Oct 2026, Accepted: 18 Oct 2026, Available online: 24 Oct 2026
©2026 The Author(s). Published by IJTLE. This is an open-access article under the CC BY license (https://creativecommons.org/licenses/by/4.0/).

Abstract

Artificial intelligence (AI) is increasingly shaping English language teaching by enabling personalized and multimodal instructional approaches. This study investigates the use of AI-customized digital picture books (DPBs) to enhance vocabulary learning in early primary classrooms. The novelty of this work lies in transforming prescribed textbook units into AI-generated, curriculum-aligned digital stories that integrate synchronized text, visuals, audio narration, and interactive features to support vocabulary acquisition. A five-week quasi-experimental design is implemented with 38 Grade 2 and Grade 3 learners following the NCERT curriculum. Vocabulary development is examined through pre- and post-assessment alongside classroom observations and teacher reflections. The findings indicate clear improvement in learners' word recognition, contextual understanding, pronunciation, and sentence usage after engagement with AI-customized DPBs. Qualitative insights further show increased learner engagement, confidence, and active participation during multimodal reading sessions. The results demonstrate that carefully designed AI-supported digital materials, when integrated with teacher mediation, create cognitively supportive and socially interactive environments for vocabulary development. This study contributes to the growing field of AI-enhanced language education by presenting an empirically grounded model for integrating generative AI tools into early ELT practice.

***Keywords*— AI in early ELT, digital picture books, multimodal learning, vocabulary development, young learners**

I. INTRODUCTION

Vocabulary Importance in Early ELT

Vocabulary knowledge is widely recognized as a central component of language proficiency and literacy development. It underpins learners' ability to comprehend texts, express meaning, and participate effectively in communication. Research consistently demonstrates that lexical knowledge is a strong predictor of reading comprehension and overall academic achievement (Nation, 2001). Furthermore, systematic and meaningful vocabulary instruction—particularly the development of high-utility academic words—plays a critical role in supporting language growth in school contexts (Beck et al., 2013).

In early English Language Teaching (ELT) classrooms, vocabulary development is especially significant. Primary learners are in a formative stage of linguistic and cognitive growth, and early exposure to meaningful vocabulary shapes later literacy outcomes. Young

learners acquire vocabulary more effectively when words are introduced through contextualized input, repetition, visual support, and interactive engagement rather than isolated memorization. However, vocabulary instruction in many early ELT settings continues to rely heavily on textbook-based word lists, which often lack sufficient multimodal support to sustain attention or promote deep understanding. Given that children are naturally responsive to stories, imagery, and sound, vocabulary teaching in primary classrooms requires more engaging and developmentally appropriate approaches. Embedding vocabulary within meaningful narrative and visual contexts can strengthen comprehension and retention.

Digital Picture Books

Digital Picture Books (DPBs) have gained attention as a practical tool for improving vocabulary instruction in primary classrooms. Unlike printed books, DPBs combine text, images, audio narration, and sometimes simple interactive features within a story format. This combination creates a richer learning experience that can capture children's interest and maintain their attention. When vocabulary is presented within stories supported by visuals and sound, learners are more likely to understand how words function in real-life contexts.

The strength of DPBs lies in their ability to connect words with images and meaningful situations. For example, when children see an illustration while hearing and reading a word, they can form stronger mental connections. Audio narration also supports pronunciation and listening skills, while visual cues help clarify meaning. Interactive elements, such as clickable words or animations, encourage active participation rather than passive reading. Instead of learning isolated vocabulary items, students encounter words within engaging narratives, which makes learning more natural and enjoyable. DPBs also support diverse learners. Visual and audio scaffolds can assist children who need additional support, while interactive elements can motivate reluctant learners. By combining multiple modes of input, DPBs create inclusive and learner-cantered environments. However, many available digital storybooks are pre-designed and may not fully match specific classroom needs or textbook content. This limitation highlights the need for more adaptable and customized digital resources.

AI-Customized Digital Picture Books

Recent developments in Artificial Intelligence (AI) offer new possibilities for designing customized Digital Picture Books that align closely with classroom objectives. AI tools allow teachers to transform traditional textbook lessons into interactive digital stories that include target vocabulary, relevant visuals, and supportive narration. Using generative AI platforms such as ChatGPT and Gemini, educators can adapt vocabulary content, simplify explanations, generate contextual sentences, and design story-based materials that suit their learners' levels.

AI customization goes beyond simply converting print materials into digital form. It allows teachers to adjust vocabulary difficulty, include culturally relevant examples, and design multimodal elements that support comprehension. For young learners, this flexibility is particularly valuable because it enables instruction to be tailored to their developmental stage and learning pace. By combining text, visuals, and audio in meaningful ways, AI-customized DPBs can help reduce confusion and strengthen memory retention. Although digital storybooks have shown positive effects in vocabulary learning, research on AI-customized DPBs in early ELT remains limited, especially at the primary school level. There is a need for empirical evidence that examines not only vocabulary improvement but also classroom experiences and teacher perspectives. Therefore, the present study explores the impact of AI-customized Digital Picture Books on vocabulary development among young learners through a structured classroom intervention, combining quantitative and qualitative data to understand their effectiveness in early ELT settings.

II. LITERATURE REVIEW

Vocabulary Acquisition and Picture Books

Vocabulary acquisition is a critical component of early literacy development, particularly in Grades 2 and 3 when learners transition from basic decoding to more complex academic reading. At this stage, lexical knowledge significantly influences comprehension, fluency, and content learning. For English language learners (ELLs), vocabulary development is especially important, as many enter school with limited exposure to academic English (Sedita, 2023). Research shows that vocabulary grows most effectively when learners encounter words repeatedly in meaningful contexts rather than through isolated memorization (Gersten et al., 2007). Picture books have long been recognized as effective tools for supporting vocabulary learning in early classrooms.

Crossref DOI: https://dx.doi.org/10.22161/ijtle.5.5.5

Children's books expose learners to richer language than everyday conversation, offering opportunities to engage with new vocabulary embedded in narrative contexts. Through illustrations and storytelling, learners infer meaning and build semantic connections (Sénéchal et al., 1995). Shared reading practices further enhance vocabulary development by providing structured discussion and guided interaction around texts (Logan et al., 2019). These practices align with Vygotsky's (1978) Sociocultural Theory, which emphasizes the importance of social interaction in learning. Vocabulary development occurs not only through exposure but also through guided discussion, questioning, and collaborative engagement. When teachers scaffold reading and discuss new words within stories, learners internalize vocabulary more effectively. Thus, picture books support vocabulary acquisition both cognitively and socially.

Digital Picture Books

With technological advancement, traditional picture books have evolved into Digital Picture Books (DPBs), which integrate multimedia features such as audio narration, animations, interactive hotspots, and contextual glosses (Li & Bus, 2023). These digital enhancements create multimodal learning environments that engage learners through visual, auditory, and interactive channels. The effectiveness of DPBs is supported by Dual Coding Theory (Paivio, 2007), which proposes that memory operates through interconnected verbal and visual systems. When vocabulary is presented simultaneously through text and corresponding imagery, learners form dual mental representations, strengthening recall. In DPBs, synchronized narration and illustrations activate both systems, enhancing vocabulary retention.

Similarly, Mayer's (2021) Cognitive Theory of Multimedia Learning suggests that learning improves when verbal and visual information are meaningfully integrated. Well-designed DPBs align text, images, and audio to reduce cognitive overload and support deeper processing. However, not all digital enhancements are beneficial. Sweller's (2019) Cognitive Load Theory warns that excessive animations or irrelevant multimedia elements can overwhelm working memory, particularly for young learners. Therefore, effective DPBs must balance engagement with cognitive clarity. Interactivity further distinguishes DPBs from traditional print books. Features such as tappable words, embedded pronunciation guides, and contextual glosses promote active engagement rather than passive reading (Tunkiel & Bus, 2022).

However, research indicates that interactivity is most effective when directly tied to instructional goals. Decorative animations that do not support vocabulary learning may distract learners and reduce comprehension (Kucirkova et al., 2021). Thus, pedagogical alignment is central to the success of digital materials. Narrative context continues to play an important role within DPBs. Stories provide repeated exposure to vocabulary within meaningful situations, allowing learners to infer meaning and strengthen retention. Emotional engagement with story characters and events further supports memory encoding (Andrä et al., 2020). In this way, DPBs extend the strengths of traditional picture books while adding multimodal and interactive dimensions.

AI in Education and Customized DPBs

The integration of Artificial Intelligence (AI) into education introduces new possibilities for customizing Digital Picture Books. Recent systematic reviews highlight the growing role of AI in enabling personalized and adaptive learning environments across educational contexts (Wang et al., 2024). AI-driven systems are increasingly used to support differentiated instruction, multimodal engagement, and intelligent feedback mechanisms that respond to learners' individual needs. Generative AI tools such as ChatGPT and Gemini allow teachers to adapt textbook content into interactive, vocabulary-focused digital narratives tailored to learners' proficiency levels. AI can generate contextual sentences, simplify complex vocabulary, and embed multimodal supports aligned with curricular goals. Research indicates that AI-supported personalization enhances learner engagement and supports instructional innovation when pedagogically guided (Younas et al., 2025).

AI customization strengthens the pedagogical value of DPBs by enabling differentiated instruction. Teachers can design stories that reflect learners' cultural contexts and adjust vocabulary difficulty according to classroom needs. From a pedagogical perspective, personalized AI systems offer opportunities for adaptive scaffolding, though effective implementation requires thoughtful instructional design and teacher mediation (Vorobyeva et al., 2025). From a cognitive perspective, AI-customized DPBs can integrate visual, auditory, and textual elements in ways that support dual coding (Paivio, 2007) and multimedia learning principles (Mayer, 2021). When designed carefully, such materials reduce unnecessary cognitive load (Sweller, 2019) and promote focused attention on essential vocabulary. However, the pedagogical effectiveness of AI tools

Crossref DOI: https://dx.doi.org/10.22161/ijtle.5.5.5

depends on structured integration rather than technological novelty alone.

Research Gap

Although previous research has demonstrated the effectiveness of picture books and digital storybooks in supporting vocabulary acquisition. No empirical studies to date have examined AI-customized DPBs in primary ELT classrooms using mixed-method designs. Existing literature often explores multimedia learning or AI applications separately, but few studies integrate cognitive theories, sociocultural perspectives, and classroom-based interventions to evaluate vocabulary outcomes among young learners. Moreover, there is a need for mixed-method research that combines quantitative measures of vocabulary improvement with qualitative insights from teacher mediation and classroom engagement. Addressing this gap, the present study investigates how AI-customized DPBs influence vocabulary development among Grade 2 and 3 learners within a structured instructional framework.

Research Questions

1. What is the effect of AI-customized Digital Picture Books (DPBs) on vocabulary development among Grade 2 and 3 learners in early ELT classrooms?
2. How do students engage with AI-customized Digital Picture Books, and what are teachers' perceptions of their effectiveness in supporting vocabulary development?

Theoretical Framework

This study is grounded in an integrated theoretical framework that combines sociocultural and cognitive perspectives to explain how AI-customized Digital Picture Books (DPBs) support vocabulary development in early ELT classrooms. Vygotsky's (1978) Sociocultural Theory highlights the importance of teacher guidance and peer interaction in facilitating the internalization of new vocabulary through scaffolded discussion and shared reading. From a cognitive perspective, Paivio's (2007) Dual Coding Theory suggests that learning is strengthened when verbal information is paired with corresponding visual representations, enabling learners to form interconnected memory pathways. Similarly, Mayer's (2021) Cognitive Theory of Multimedia Learning emphasizes that meaningful integration of text, visuals, and audio enhances comprehension by supporting dual-channel processing. Complementing these principles, Sweller's (2019) Cognitive Load Theory underscores the need to design digital materials that minimize extraneous cognitive demands so learners can focus on essential vocabulary. Together, these theoretical perspectives provide a coherent foundation for examining how well-designed, teacher-mediated AI-customized DPBs can enhance vocabulary acquisition among young learners in early ELT contexts. These theoretical perspectives collectively justify the integration of AI-customized DPBs as a cognitively and socially mediated instructional approach.

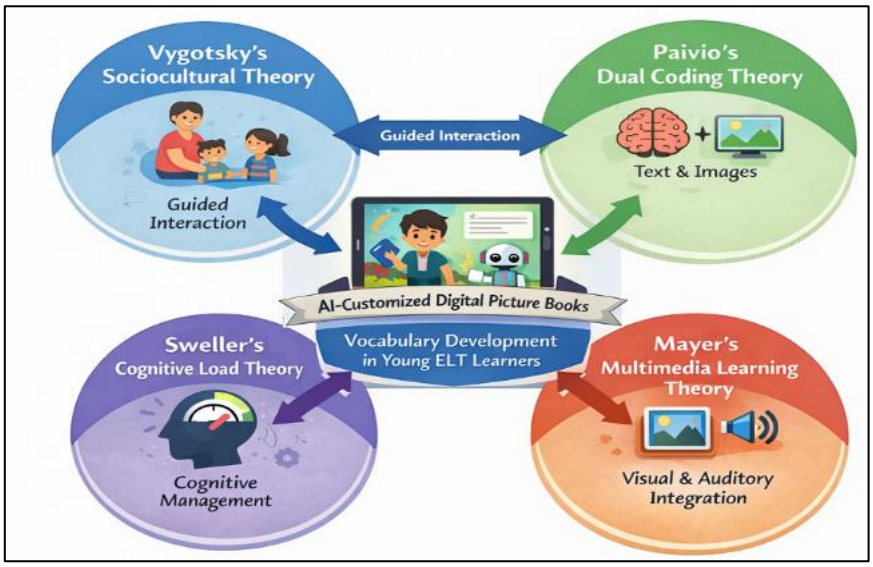


Fig.1. Theoretical Model for AI-Customized DPBs

The above figure illustrates the integrated theoretical foundation of the study. Vygotsky's Sociocultural Theory emphasizes guided interaction and teacher mediation in vocabulary learning. Paivio's Dual Coding

Crossref DOI: https://dx.doi.org/10.22161/ijtle.5.5.5

Theory highlights the integration of verbal and visual systems in memory formation. Mayer's Cognitive Theory of Multimedia Learning supports the meaningful alignment of text, visuals, and audio to enhance comprehension. Sweller's Cognitive Load Theory underscores the importance of reducing extraneous cognitive demands in multimedia design. Together, these theories inform the use of AI-customized Digital Picture Books (DPBs) to support vocabulary development among young learners in early ELT classrooms.

III. RESEARCH METHODOLOGY

Research Design

This study adopted a mixed-method quasi-experimental design to examine the effect of AI-customized Digital Picture Books (DPBs) on vocabulary development in early ELT classrooms. A pre-test–post-test approach was employed to measure changes in students' vocabulary knowledge following a five-week instructional intervention. Quantitative data were collected through structured vocabulary assessments, while qualitative data were gathered through classroom observations and teacher reflections to provide additional insight into learner engagement and instructional effectiveness.

Participants

The participants were 38 students enrolled in Grades 2 and 3 at a private primary school following the NCERT curriculum. The students, aged approximately 7–9 years, were learning English as part of their regular school program and represented mixed proficiency levels typical of early ELT classrooms. In addition, five English language teachers contributed qualitative data through structured reflections and classroom observations. Necessary permissions were obtained from the school administration, and parental consent was secured prior to the study.

Procedure

The study was conducted over five weeks during regular English instructional plan detail plan find table.1. In the first week, students completed a vocabulary pre-test to establish baseline knowledge. Selected units from the prescribed textbook were then transformed into AI-customized Digital Picture Books using generative AI tools such as ChatGPT and Gemini. During the intervention, teachers implemented the DPBs through guided reading, interactive discussion, and vocabulary-focused activities. Students engaged with multimodal features such as synchronized text, visuals, and audio support. Teacher mediation was incorporated through questioning, clarification of meanings, and reinforcement activities. At the end of the five-week intervention, students completed a post-test to measure vocabulary gains. A paired-sample t-test was conducted.

Table 1. Five-Week Intervention Plan

Week	Focus	Activities	Teacher Role
Week 1	Pre-test & Introduction	Conduct pre-test; introduce AI-customized DPBs; demonstrate text, visuals, and audio features	Explain objectives; model guided reading
Week 2	Visual & Contextual Learning	Guided reading of DPB Unit 1; picture–word matching; vocabulary discussion	Scaffold meaning through visuals; ask comprehension questions
Week 3	Interactivity & Pronunciation	DPB Unit 2; clickable glosses; audio pronunciation practice; pair activities	Provide pronunciation support; monitor engagement
Week 4	Contextual Usage	Sentence formation; short storytelling using target words; mini digital tasks	Support vocabulary uses in sentences; guide storytelling
Week 5	Review & Post-test	Vocabulary recap through DPBs; interactive revision; conduct post-test	Reinforce learning; administer assessment

Pre-Test

A vocabulary pre-test was conducted to assess students' baseline knowledge of selected words from Grade 2, Unit 1: *Good Morning* (Poem) and Grade 3, Unit 2: *Seeing Without Seeing*. A total of 25 target vocabulary items were selected from these units and later incorporated into the AI-customized Digital Picture Books used during the intervention. The test consisted of five task types: picture–word matching, multiple-choice meaning recognition, fill-in-the-blanks, sentence construction, and oral word identification. Each section contained five items, with one mark awarded per correct response,

Crossref DOI: https://dx.doi.org/10.22161/ijtle.5.5.5

resulting in a total score of 25 marks. The assessment measured both receptive and productive vocabulary knowledge.

The mean pre-test score of the 38 learners was 10.4 (SD = 2.8), with a median score of 10, indicating limited to moderate baseline vocabulary knowledge. Most scores fell within the low to medium range (6–12 marks), reflecting developing vocabulary proficiency prior to the intervention. The vocabulary development rubric was developed for this study based on established principles of analytic rubric design and formative assessment (Andrade, 2000; Brookhart, 2013)

Table 2. Vocabulary Development Assessment Rubric

Criteria	Excellent	Good	Developing	Beginning
Word Recognition	Identifies all words correctly	Few errors	Recognizes some words	Struggles to identify words
Meaning	Clear understanding in context	Minor support needed	Partial understanding	Limited understanding
Usage	Uses words correctly in sentences	Minor errors	Limited sentence use	Unable to use words properly
Pronunciation	Clear and accurate	Minor errors	Noticeable difficulty	Incorrect pronunciation
Engagement	Actively participates	Participates regularly	Needs prompting	Passive participation



Fig.2. Sample AI-Customized DPBs Pages from Grade 2 (Left) and Grade 3 (Right)

Post-Test

The vocabulary development rubric had a maximum score of 20 marks, derived from five criteria evaluated on a four-point scale. Students' total scores were categorized into four performance levels to interpret vocabulary growth. Scores ranging from 14–19 were classified as High, indicating strong vocabulary development; 9–14 as Medium, reflecting moderate proficiency; 5–9 as Low, indicating limited vocabulary control; and 1–5 as Needs Support, suggesting minimal vocabulary knowledge requiring significant instructional assistance. This classification helped identify learners' proficiency levels before and after the AI-customized Digital Picture Book intervention and supported comparative analysis of vocabulary improvement.

Following the five-week AI-customized Digital Picture Book (DPB) intervention, a post-test was administered using the same format and scoring criteria as the pre-test. The assessment measured students' receptive and productive vocabulary knowledge, including word

Crossref DOI: https://dx.doi.org/10.22161/ijtle.5.5.5

recognition, contextual understanding, sentence usage, and pronunciation.

The results demonstrated notable improvement in vocabulary performance. The mean post-test score increased to 16.8 (SD = 2.4), compared to the pre-test mean of 10.4 (SD = 2.8). The median post-test score was 17, indicating a clear shift toward higher vocabulary proficiency. Most students scored within the high-performance range (14–19 marks), reflecting substantial vocabulary growth following the intervention. The reduction in standard deviation further suggests improved consistency in learners' performance. These findings indicate that the AI-customized DPB intervention positively influenced vocabulary development among Grade 2 and 3 learners.

Table 3. Comparison of Pre-Test and Post-Test

Test	Mean	SD	Median	Performance Range
Pre-Test	10.4	2.8	10	6–12 (Low–Medium)
Post-Test	16.8	2.4	17	14–19 (High)

Table 3 presents the descriptive statistics comparing students' vocabulary performance before and after the AI-customized Digital Picture Book (DPB) intervention. The mean pre-test score was **10.4 (SD = 2.8)**, with a median of **10**, reflecting low to moderate baseline vocabulary knowledge. Following the five-week intervention, the mean post-test score increased to **16.8 (SD = 2.4)**, with a median of **17**, indicating a clear improvement in vocabulary proficiency. The majority of students shifted from the low–medium performance range (6–12 marks) to the high range (14–19 marks).

Statistical analysis indicated that the increase in vocabulary scores was **significant (p < .001)**, demonstrating that the improvement observed after the intervention was unlikely due to chance. The reduction in standard deviation in the post-test further suggests greater consistency in students' performance. Overall, the findings indicate that the AI-customized DPB intervention contributed to meaningful vocabulary development among Grade 2 and Grade 3 learners.

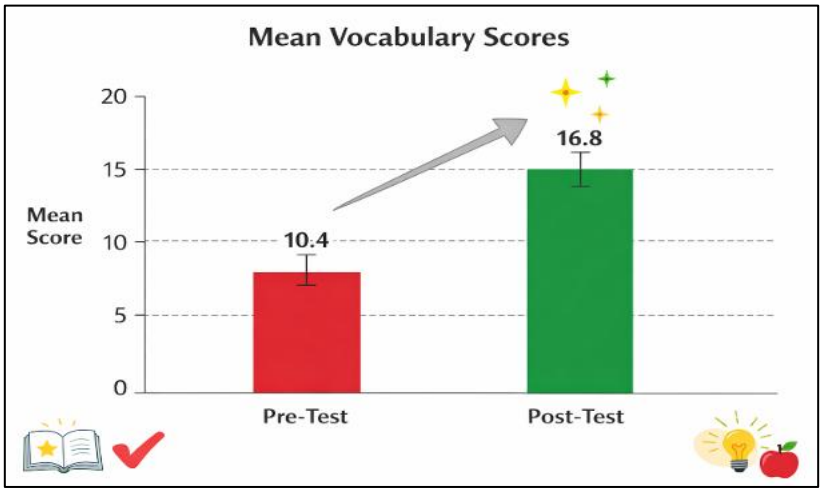


Fig.3. Mean Scores of Pre- and Post- Test

The figure illustrates the comparison of mean vocabulary scores before and after the five-week AI-customized Digital Picture Book (DPB) intervention. The red bar represents the pre-test mean score (10.4), while the green bar represents the post-test mean score (16.8). The visual comparison shows an increase of **6.4** in the mean vocabulary score following the intervention. Horizontal grid lines enhance clarity and allow for easier interpretation of score differences. This upward shift visually reinforces the statistical findings (p < .001), demonstrating that the five-week DPB intervention contributed significantly to vocabulary development among Grade 2 and Grade 3 learners.

Data Analysis

Crossref DOI: https://dx.doi.org/10.22161/ijtle.5.5.5

The mixed-method analysis revealed substantial improvement in students' vocabulary performance following the five-week AI-customized Digital Picture Book (DPB) intervention. Quantitative findings indicated that the mean pre-test score of 10.4 (SD = 2.8) increased to 16.8 (SD = 2.4) in the post-test, reflecting a mean gain of 6.4 points. Statistical analysis confirmed that this improvement was significant ($p < .001$), indicating that the vocabulary gains were unlikely due to chance. A clear shift was observed from the low–medium performance range to the high range after the intervention, demonstrating meaningful vocabulary growth among Grade 2 and Grade 3 learners.

Classroom observations provided further evidence of instructional impact. During DPB sessions, students appeared more attentive and actively engaged compared to traditional textbook lessons. They responded enthusiastically to interactive features, particularly audio pronunciation tools and visual glosses. Repeated exposure to vocabulary within narrative contexts enabled learners to infer meanings more confidently and use target words during discussions, supporting findings that contextualized and multimodal vocabulary presentation enhances retention and comprehension (Paivio, 2007; Mayer, 2021).

Qualitative insights were derived from classroom observations and teacher reflections during the five-week intervention. Observations indicated that students were noticeably more attentive and engaged during AI-customized Digital Picture Book (DPB) sessions compared to conventional textbook-based instruction. Learners interacted actively with multimodal features, including visuals, synchronized text, and audio narration, which appeared to support vocabulary comprehension and recall. Increased participation during guided discussions further suggested improved learner confidence and motivation, consistent with sociocultural perspectives emphasizing guided interaction in vocabulary development (Vygotsky, 1978).

T1 *noted, "Students were more confident in answering vocabulary questions when they saw the pictures along with the words."* This reflection reflects the theme of visual scaffolding and confidence building, indicating that the integration of images with text strengthened comprehension and reduced hesitation in vocabulary use, aligning with dual coding principles (Paivio, 2007).

T3 *reflected, "Through the DPBs, children understood the meaning better instead of just memorizing words."* This comment highlights the theme of contextual and narrative-based learning, suggesting that embedding vocabulary within meaningful story contexts promoted deeper semantic understanding rather than surface-level memorization, consistent with multimedia learning theory (Mayer, 2021).

T4 *observed, "Interactive features engaged students focused, and they repeated pronunciation without fear."* This observation represents the theme of interactive engagement and pronunciation reinforcement, demonstrating that multimodal elements encouraged repeated exposure and active participation in vocabulary practice while maintaining manageable cognitive demands (Sweller, 2019).

The thematic analysis reveals three dominant dimensions: visual reinforcement, narrative contextualization, and interactive engagement. These interconnected themes help explain the quantitative gains observed in the post-test results. The integration of multimodal design with teacher-guided mediation appears to have created a supportive and cognitively balanced learning environment that enhanced both receptive and productive vocabulary development among young learners in early ELT classrooms.

Ethical Considerations

Permission to conduct this study was obtained from the head of the participating institution. Informed consent was also taken from the parents or guardians of all student participants before the study began. Teachers who shared their reflections participated voluntarily and were informed about the purpose of the research.

To protect privacy and confidentiality, the name of the school has not been mentioned in this paper. Individual students' marks and personal details have not been disclosed. All data are presented in summarized form, and teacher responses are identified using codes (e.g., T1, T3, T4). The study followed ethical guidelines for research involving human participants and ensured that no student was disadvantaged academically or personally during the intervention.

IV. CONCLUSION

This study examined the impact of AI-customized Digital Picture Books (DPBs) on vocabulary development among Grade 2 and Grade 3 learners in early ELT classrooms. The findings demonstrated that integrating multimodal digital resources into structured classroom instruction can meaningfully enhance vocabulary acquisition. Students showed clear improvement in

Crossref DOI: https://dx.doi.org/10.22161/ijtle.5.5.5

word recognition, contextual understanding, pronunciation, and sentence usage following the five-week intervention.

Beyond measurable gains, qualitative insights revealed increased learner engagement, confidence, and participation during vocabulary activities. Visual scaffolding, narrative embedding of target words, and interactive pronunciation features contributed to deeper semantic understanding rather than rote memorization. These outcomes align with principles of dual coding and multimedia learning, which emphasize the integration of verbal and visual input to strengthen retention and comprehension (Paivio, 2007; Mayer, 2021).

Teacher mediation further reinforced vocabulary retention through guided discussion and contextual practice, consistent with sociocultural perspectives on scaffolded learning (Vygotsky, 1978).

Overall, the study underscores the pedagogical value of AI-supported Digital Picture Books as an innovative and practical approach to vocabulary instruction in early ELT settings. When thoughtfully designed to minimize cognitive overload and maximize meaningful engagement (Sweller, 2019), AI-customized DPBs can create cognitively supportive learning environments that promote sustained vocabulary growth. More broadly, the findings suggest that carefully integrated AI tools can complement teacher-led instruction and contribute to the evolving landscape of technology-enhanced language education in primary classrooms.

Limitations

Despite its positive findings, the study has several limitations. First, the use of free versions of AI tools limited the ability to convert entire textbook units simultaneously; content had to be transformed page by page, which was time-consuming and may affect scalability in larger implementations. Second, the study was conducted within a single private school context, which may limit generalizability to other educational settings, particularly government schools with differing technological resources.

Additionally, teachers had limited prior training in AI-integrated instructional design. While orientation was provided, more comprehensive professional development could enhance effective implementation. Technological infrastructure, including stable internet access and availability of digital devices, also presents practical challenges for broader adoption. Furthermore, the absence of a control group and the relatively short duration of the intervention restrict the ability to make broader causal claims or assess long-term vocabulary retention. Future research may address these limitations by expanding to diverse school contexts, incorporating structured teacher training programs, and examining longitudinal vocabulary outcomes.

Suggestions

Based on the findings of this study, it is suggested that AI-customized Digital Picture Books (DPBs) be incorporated as a supplementary strategy for vocabulary instruction in early ELT classrooms. Teachers should integrate multimodal features—such as visuals, audio narration, and interactive vocabulary supports—alongside guided discussion to enhance comprehension and retention. Effective implementation requires pedagogical alignment and purposeful teacher mediation rather than reliance on technological novelty alone (Wang et al., 2024; Younas et al., 2025). Structured professional development programs are recommended to strengthen teachers' ability to design and implement AI-supported instructional materials effectively.

Schools and policymakers should also consider improving digital infrastructure to facilitate sustainable technology integration. For future research, longer-term studies are recommended to examine vocabulary retention beyond short-term interventions. Comparative studies across diverse educational settings, including government and rural schools, would enhance generalizability. Further investigation into the impact of AI-customized DPBs on other language skills, such as reading comprehension and speaking fluency, would contribute to a broader understanding of AI-supported multimodal learning in early ELT contexts**.**

Data Availability

The data supporting the findings of this study are available from the corresponding author upon reasonable request. Due to ethical considerations and institutional policies protecting the privacy of participating students and teachers, individual-level data cannot be made publicly available. All reported results are presented in aggregated form to ensure confidentiality.

CONFLICTS OF INTEREST

The author declares that there is no conflict of interest regarding the publication of this paper.

FUNDING STATEMENT

This research received no external funding.

AUTHORS' CONTRIBUTIONS

Conceptualization, S.B.; Methodology, S.B.; Software, S.B.; Validation, S.B.; Formal Analysis, S.B.; Investigation, S.B.; Resources, S.B.; Data Curation, S.B.; Writing – Original Draft Preparation, S.B.; Writing – Review & Editing, S.B.; Visualization, S.B.; Supervision, S.B.; Project Administration, S.B.; Funding Acquisition, Not applicable.

ACKNOWLEDGMENT

The author expresses sincere gratitude to the participating school, teachers, students, and parents for their cooperation and support throughout the study. Their collaboration made this research possible

REFERENCES

[1] Andrä, C., Mathias, B., Schwager, A., Macedonia, M., & von Kriegstein, K. (2020). Learning foreign language vocabulary with gestures and pictures enhances memory. *Scientific Reports, 10*(1), Article 1. https://doi.org/10.1038/s41598-020-61852-z

[2] Andrade, H. G. (2000). Using rubrics to promote thinking and learning. *Educational Leadership, 57*(5), 13–18.

[3] Beck, I. L., McKeown, M. G., & Kucan, L. (2013). *Bringing words to life: Robust vocabulary instruction* (2nd ed.). Guilford Press.

[4] Brookhart, S. M. (2013). *How to create and use rubrics for formative assessment and grading*. ASCD.

[5] Gersten, R., Fuchs, L. S., Williams, J. P., & Baker, S. (2007). Teaching reading comprehension strategies to students with learning disabilities: A review of research. *Review of Educational Research, 77*(3), 387–437. https://doi.org/10.3102/0034654307302176

[6] Kucirkova, N., Furenes, M. I., & Bus, A. G. (2021). The role of digital storybooks in supporting young children's vocabulary development: A meta-analysis. *Review of Educational Research, 91*(6), 851–885. https://doi.org/10.3102/00346543211046739

[7] Li, X., & Bus, A. G. (2023). Multimedia storybooks and vocabulary learning: A systematic review and meta-analysis. *Educational Psychology Review, 35*(2), 1–28. https://doi.org/10.1007/s10648-023-09727-3

[8] Logan, J. A. R., Justice, L. M., Yumus, M., & Chaparro-Moreno, L. J. (2019). When children are not read to at home: The million word gap. *Journal of Developmental & Behavioral Pediatrics, 40*(5), 383–386. https://doi.org/10.1097/DBP.0000000000000657

[9] Mayer, R. E. (2021). *Multimedia learning* (3rd ed.). Cambridge University Press.

[10] Moskal, B. M. (2000). Scoring rubrics: What, when and how? *Practical Assessment, Research & Evaluation, 7*(3), 1–7. https://doi.org/10.7275/a5vq-7q66

[11] Nation, I. S. P. (2001). *Learning vocabulary in another language*. Cambridge University Press.

[12] Paivio, A. (2007). *Mind and its evolution: A dual coding theoretical approach*. Lawrence Erlbaum Associates.

[13] Sedita, J. (2023). Vocabulary instruction: Key research and implications for practice. *Perspectives on Language and Literacy, 49*(1), 12–17.

[14] Sénéchal, M., Thomas, E., & Monker, J. (1995). Individual differences in 4-year-old children's acquisition of vocabulary during storybook reading. *Journal of Educational Psychology, 87*(2), 218–229.

[15] Sweller, J. (2019). Cognitive load theory and educational technology. *Educational Technology Research and Development, 67*(1), 1–16. https://doi.org/10.1007/s11423-018-9615-4

[16] Tunkiel, K., & Bus, A. G. (2022). The impact of interactivity in digital storybooks on young children's vocabulary development. *Computers & Education, 186*, Article 104530. https://doi.org/10.1016/j.compedu.2022.104530

[17] Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Harvard University Press.

[18] Vorobyeva, K. I., Belous, S., Savchenko, N. V., Smirnova, L. M., Nikitina, S. A., & Zhdanov, S. P. (2025). Personalized learning through AI: Pedagogical approaches and critical insights. *Contemporary Educational Technology, 17*(2), ep574. https://doi.org/10.30935/cedtech/16108

[19] Wang, S., Wang, F., Zhu, Z., Wang, J., Tran, T., & Du, Z. (2024). Artificial intelligence in education: A systematic literature review. *Expert Systems with Applications, 232*, 124167. https://doi.org/10.1016/j.eswa.2024.124167

[20] Younas, M. Y., El-Dakhs, D. A. S., & Noor, U. (2025). The impact of artificial intelligence-based learning tools in academic innovation: A review. *Frontiers in Education*, 1689205. https://doi.org/10.3389/feduc.2025.1689205

Author's Bio

Banoth Suresh is a Ph.D. Research Scholar in English Language Education at the English and Foreign Languages University (EFLU), Hyderabad, India. His primary research interests include artificial intelligence (AI)-based tools in English Language Teaching (ELT), technology-enhanced language learning, and vocabulary development among young and secondary school learners. He is particularly interested in exploring the use of AI-powered educational tools, digital storytelling, and digital picture books to enhance vocabulary acquisition, learner engagement, and English language proficiency. His research focuses on the pedagogical integration of emerging technologies to develop innovative, interactive, and learner-centred approaches to English language teaching and learning.

Dr. N. Narendar is the Principal of Vidyarthi College of Education, Mavala, Adilabad, Telangana, India. He holds a Ph.D.in Education from Kalinga University and is engaged in teacher education and academic leadership. His professional interests include teacher education, educational practices, curriculum development, and the professional development of teachers. Through his leadership and academic engagement, he contributes to the advancement of teacher education and the improvement of educational practices.

International Journal of Teaching, Learning and Education (IJTLE)

"Advancing Research in Teaching, Learning and Education Globally"

The International Journal of Teaching, Learning and Education (IJTLE) is an international, open-access, double-blind peer-reviewed journal committed to publishing high-quality research in education, language, and interdisciplinary studies.

Aim & Scope

IJTLE focuses on teaching methodologies, learning processes, curriculum development, educational policy, language studies, and interdisciplinary research.

Key Highlights

Open Access Journal

Double-Blind Peer Review

Bi-monthly Publication

International Editorial Board

DOI for all papers

Types of Papers

Research Articles, Review Papers, Case Studies, Conceptual Papers, Short Communications

Call for Papers

IJTLE invites original and unpublished manuscripts for upcoming issues.

Contact & Submission

Website: https://ijtle.com

Email: editor.ijtle@gmail.com

ISSN: 2583-4371

Frequency: Bi-monthly

Copyright © International Journal of Teaching, Learning and Education