



Bridging the Modality Gap: How Digital Visualisation Supports Syntactic Scaffolding in English Reading Comprehension among Tanzanian Primary Learners

Alpha Kavishe, Kulwa Matalu & Andwele Mwakasege

University of Dodoma, Tanzania

Article History

Received: 02.07.2026

Revised: 19.09.2026

Accepted: 24.09.2026

Published: 26.09.2026

Keywords

Comprehension

Digital visualisation

Reading

Scaffolding

How to cite:

Kavishe, A., Matalu, K., & Mwakasege, A. (2026). Bridging the Modality Gap: How Digital Visualisation Supports Syntactic Scaffolding in English Reading Comprehension among Tanzanian Primary Learners. *Journal of Linguistics, Literary and Communication Studies*, 5(2), 34-47.

Abstract

Tanzanian primary school learners face persistent challenges in English reading comprehension, particularly in processing complex syntactic structures that differ substantially from their native Kiswahili. This syntactic bottleneck limits their ability to progress beyond literal comprehension to inferential and critical reading. Despite the promise of digital visualisation, the specific mechanisms through which it supports comprehension remain undertheorised, and existing research has not examined differential effects across comprehension levels in African educational contexts. This study investigates how dynamic digital visualisation supports English reading comprehension among these learners. It examines differential effects across literal, inferential, and critical comprehension levels and proposes a Visual-Syntactic Scaffolding (VSS) Model as an inferential framework. A mixed-methods case study was conducted in ten purposively selected primary schools in Morogoro Municipality, Tanzania, involving 100 Standard VI pupils and 90 English language teachers. Purposive sampling was adopted because the intervention required schools with reliable electricity, basic digital devices, and willingness to participate in a 20-week intervention. The 20-week Playcast intervention combined dynamic video, colour-coded syntactic formatting, interactive features, and embedded comprehension questions, with structured teacher training and ongoing support. Data were collected through pre- and post-intervention comprehension assessments, teacher surveys, classroom observations, and semi-structured interviews. Quantitative data were analysed using paired *t*-tests and effect sizes; qualitative data were analysed thematically. The intervention produced significant overall comprehension gains ($d^* = 0.81$), with strongest effects at literal ($d^* = 0.87$), moderate at inferential ($d^* = 0.54$), and weakest at critical ($d^* = 0.22$) levels. Teacher data indicate high intervention acceptability, pedagogical feasibility, and implementation fidelity. Lower proficiency students relied on visual cues for meaning construction; higher proficiency students used visualisation for verification. The study

Copyright © 2026





measures global comprehension outcomes rather than real-time psycholinguistic processing; the VSS Model is presented as an inferential framework requiring validation through experimental designs with eye-tracking and think-aloud protocols. The VSS Model explains how dynamic visualisation may support comprehension through three mechanisms: syntactic bootstrapping, cognitive load reduction, and external mental model construction.

Introduction

Tanzanian primary school learners face a well-documented reading comprehension crisis. Most pupils struggle to understand main ideas, summarise texts, and read English books independently (Mwanga & Mwanga, 2026). This crisis is driven in part by a syntactic bottleneck created by the structural distance between Kiswahili and English (Tibategeza & Plessis, 2021). Kiswahili, an agglutinative Bantu language with a noun-class system and relatively flexible word order, differs substantially from English, an analytic West Germanic language whose canonical Subject-Verb-Object (SVO) order undergoes systematic transformation through passivisation, inversion, interrogation, and clefting, alongside complex tense-aspect morphology. These structural differences impose heavy cognitive demands on parsing English sentence structures, leaving insufficient cognitive resources for higher-order comprehension.

Reading comprehension in a second language (L2) presents challenges that differ from those faced by first language (L1) readers. While both L1 and L2 readers must navigate the interplay between decoding, syntactic processing, and semantic integration, L2 readers do so with an underdeveloped linguistic baseline in the target language and must contend with cross-linguistic interference from their native tongue (Bordag et al., 2025). For Tanzanian primary school learners, who acquire English as a second language within a multilingual educational context where home languages, Kiswahili, and English interact, these challenges are particularly acute.

Digital visualisation has emerged as a promising intervention to support L2 reading comprehension across multiple educational contexts (Asencios & Jaramillo, 2025). Recent research has focused specifically on Visual-Syntactic Text Formatting (VSTF), a technique that transforms block-shaped text into cascading patterns that reflect syntactic structure, making grammatical relationships visually explicit (Tate et al., 2019). Empirical studies have demonstrated that VSTF-based text can significantly improve reading speed, comprehension, and retention, particularly among lower proficiency learners (Ozaki & Ueda, 2020; Park et al., 2016; Gao et al., 2022). Static VSTF, however, presents syntactic structure in a fixed, non-interactive format. This study extends static VSTF into a dynamic multimedia environment that combines cascading syntactic formatting with video content, making both syntactic and event relationships visually explicit.

In the Tanzanian context, research has documented both the need for reading comprehension interventions and the challenges of implementing them. Kavishe et al. (2026) found that teachers demonstrate low competence in using digital visualisation materials, with contributing factors including negative attitudes, low self-efficacy, limited access to technological resources, inadequate professional training, and lack of institutional support. The Tanzania Institute of Education (2020) has developed curriculum guidelines emphasising interactive and visual strategies for reading instruction, yet implementation remains constrained by resource limitations (Ngonyani et al., 2026).



This paper addresses the gap between established pedagogical effectiveness and theoretical understanding by examining how dynamic Playcast-based visualisation supports reading comprehension. The study examines the differential effects of visualisation across literal, inferential, and critical levels of comprehension, investigating whether the observed pattern of effects is consistent with syntactic scaffolding as an explanatory framework.

Significance of the Study

This study makes theoretical, pedagogical, and policy contributions. Theoretically, it extends research on Visual-Syntactic Text Formatting to the African educational context and examines whether the observed pattern of comprehension gains is consistent with the syntactic bottleneck hypothesis (Ozaki & Ueda, 2020). By investigating differential effects across comprehension levels, the study contributes to a more nuanced understanding of what visualisation can and cannot achieve in L2 reading (Atanasio Asencios & Saavedra Jaramillo, 2025). Pedagogically, the findings provide guidance for the design of digital reading comprehension materials. Practically, the study informs educational technology policy by examining implementation challenges including limited access to computers, internet connectivity, and reliable electrical power (Ngonyani et al., 2026).

Literature Review

The Nature of L2 Reading Comprehension

Reading comprehension is a complex cognitive process involving the integration of multiple component skills. The Simple View of Reading provides a foundational framework: reading comprehension is the product of decoding (word recognition) and linguistic comprehension (language comprehension). Recent validation studies have confirmed the applicability of this framework across diverse learner populations, including L2 readers in varied educational contexts, with findings indicating that language comprehension becomes increasingly predictive of reading comprehension as learners progress beyond basic decoding stages (Nordström et al., 2025). Once words are successfully decoded, students need to use their underlying oral language comprehension to understand what is being communicated.

For L2 readers, the relative contribution of these components differs from L1 readers. Research has consistently shown that for many L2 learners, particularly those in educational contexts where L2 is acquired primarily through formal instruction, the primary challenge lies not in word decoding but in language comprehension (Park et al., 2012). This suggests that the bottleneck in L2 reading is located at the level of linguistic processing rather than basic decoding.

The comprehension process can be conceptualised in terms of three interacting levels: lexical access (word recognition), syntactic parsing (grammatical analysis), and semantic integration (construction of meaning). For L2 readers, syntactic parsing is frequently the most demanding component, particularly when the L1 and L2 differ substantially in syntactic structure. As Park et al. (2012) note, the primary challenge for L2 readers lies not in word decoding but in language comprehension, suggesting that syntactic scaffolding may be more effective than vocabulary support alone.

Research conducted in the Tanzanian context highlights the severity of these challenges. Mwanga and Mwanga (2026) found that most pupils struggled with understanding main ideas, summarising, and reading English books. Further research by Ngonyani et al. (2026) identified multiple barriers to effective reading comprehension instruction, including insufficient reading materials, limited access to computers, overhead projectors, televisions, internet connectivity, and reliable electrical power.



Levels of Reading Comprehension

Reading comprehension is commonly classified into hierarchical levels, with the most widely recognised taxonomy being Barrett's (1976), which distinguishes literal, inferential, evaluative, and appreciative comprehension. This taxonomy continues to be widely applied in contemporary reading research; for example, Chotimah et al. (2026) used Barrett's framework to develop a LOTS and HOTS-based reading comprehension instrument for lower secondary students. A recent methodological study also integrated Barrett's taxonomy into a digital platform for developing critical reading skills, demonstrating the taxonomy's continued relevance for technology-enhanced reading instruction (Şimşek & Direkci, 2023). Literal comprehension involves understanding information explicitly stated in the text. For L2 readers, this requires accurate lexical access and syntactic parsing to identify agents, actions, and objects. Inferential comprehension requires the reader to go beyond the explicit text, making connections between ideas and filling in missing information. This process draws heavily on pragmatic processing, including the resolution of anaphora and the drawing of bridging inferences. Evaluative and critical comprehension demand metalinguistic awareness, allowing the reader to assess the author's intent, evaluate arguments, and consider multiple perspectives. For L2 readers, syntactic processing demands may prevent progression beyond literal comprehension.

Digital Visualisation and L2 Reading

Digital visualisation encompasses a range of techniques for representing textual information through visual means, including static images, animations, video, and interactive multimedia (Masika et al. 2023). The role of visualisation in reading comprehension has been examined through several theoretical lenses. Dual Coding Theory (Paivio, 1986) suggests that information processed through both verbal and visual channels is better remembered and understood than information processed through a single channel. Recent research on L2 vocabulary acquisition continues to draw on dual coding principles, demonstrating that verbal representations paired with visual imagery yield more consistent results in boosting short- and long-term recall than single-channel encoding (Zhang et al., 2025). Multimedia learning theory extends this framework, proposing that well-designed multimedia materials reduce cognitive load and facilitate learning (Mayer, 2005).

Recent research has focused on Visual-Syntactic Text Formatting (VSTF), a technique that transforms block-shaped text into cascading patterns that help readers identify grammatical structure (Tate et al., 2019). VSTF presents texts in a cascaded version that compensates for limitations in explicit or implicit syntactic knowledge among second language learners (Park et al., 2012). By visually grouping words into syntactic constituents, VSTF reduces the working memory demands of syntactic parsing. Park et al. (2016) investigated the effects of VSTF on Korean middle school EFL learners' reading comprehension over 15 weeks, finding that VSTF groups significantly outperformed controls on reading comprehension ($*p < .01$). Ozaki and Ueda (2020) found that low proficiency Japanese high school students showed significant increases in reading speed, comprehension, reading efficiency, and retention when using VSTF-based text. Walker et al. (2005) confirmed an immediate 40 percent improvement in reading comprehension and a significant reduction in eye fatigue when using VSTF.

The Case for Syntactic Scaffolding

The body of research on L2 reading and digital visualisation supports a specific focus on syntactic scaffolding. Park (2017) reviewed technological tools for syntactic enhancement, arguing that VSTF technology may help increase syntactic awareness. The promise of syntactic scaffolding lies in its potential to address the primary bottleneck in L2 reading: the processing of unfamiliar syntactic structures. While multimedia scaffolding receives broader attention, empirical evidence highlights the promising value of syntactic scaffolding, particularly through VSTF (Park et al., 2012).



Theoretical Framework

This study is grounded in two complementary theoretical perspectives. Relevance Theory (Sperber & Wilson, 1995) explains that human cognition seeks maximum cognitive effects with minimum processing effort. For L2 readers, syntactic complexity increases processing effort, leaving insufficient capacity for semantic integration and inference. Digital visualisation reduces this effort by making syntactic relationships visually explicit. This aligns with the broader principle of external scaffolding, where learners use external representations to support internal cognitive processes. If visualisation primarily reduces syntactic processing effort, its effects should be most pronounced at the literal level, where syntactic parsing is the primary bottleneck.

Visual-Syntactic Text Formatting (VSTF) reformats block-shaped text into cascading patterns reflecting syntactic structure, making grammatical relationships visually explicit (Tate et al., 2019) and reducing working memory demands of syntactic parsing (Park et al., 2012). The Playcast-based visualisation used in Tanzania shares VSTF features while extending them through dynamic video content that makes temporal and causal relationships visually explicit. Relevance Theory and VSTF are complementary: Relevance Theory provides the broader cognitive lens for why reducing processing effort matters, while VSTF provides the specific mechanism. Together they predict that visualisation will reduce processing effort, free cognitive resources for meaning construction, show strongest effects at the literal level, and diminish at higher levels where pragmatic and metacognitive processes become more important.

The Interactive-Compensatory Model (Stanovich, 1980, as cited in Kosaka, 2025) provides an additional lens for understanding why proficiency levels interact differently with visuals. This model suggests that lower proficiency readers rely more heavily on compensatory strategies, including visual cues, when syntactic knowledge is limited. Higher proficiency readers, by contrast, have sufficient syntactic knowledge to parse text independently and use visual cues primarily for verification.

Method

Research Design

This study employed a mixed-methods case study design, following the approach established in previous research in the Tanzanian context (Kavishe et al., 2026; Mwanga & Mwanga, 2026). The study was conducted in ten selected primary schools in Morogoro Municipality, Tanzania. The study focused on Standard VI pupils, who have acquired basic reading skills and are transitioning to more complex reading comprehension tasks.

Sampling and Sampling Method

A purposive sampling strategy was adopted to select ten primary schools in Morogoro Municipality. Purposive sampling was appropriate because the study required schools with: (a) access to electricity and basic digital devices (tablets or projectors); (b) willingness to participate in a 20-week intervention; and (c) Standard VI classes with English language instruction. From each school, one Standard VI class was selected, yielding a total of 100 pupils (approximately 10 pupils per school). The pupil sample size was determined by the number of pupils enrolled in the selected Standard VI classes and the availability of consent from parents and assent from pupils. All 100 pupils in the selected classes participated in the intervention and completed both pre- and post-intervention assessments.

From each of the ten schools, all English language teachers who taught Standard VI were invited to participate. This yielded 90 English language teachers across the ten schools (ranging from 7 to 12



teachers per school). In addition, all 10 head teachers from the participating schools were included because they provided institutional perspectives on implementation. The teacher sample size was therefore determined by the total number of English language teachers and head teachers in the participating schools, not by a priori power calculation.

The two strata (pupils and teachers) were necessary because they served different research purposes: pupils were the primary unit of analysis for measuring comprehension outcomes, while teachers provided data on intervention acceptability, pedagogical feasibility, and implementation fidelity.

The Playcast Intervention

The Playcast intervention involved the use of dynamic, video-based visualisation materials for reading comprehension instruction. Playcast materials were developed to support reading comprehension of English texts, incorporating: (1) dynamic video content showing actions and events described in the text; (2) textual presentation with colour coding and spatial formatting to support syntactic parsing; (3) interactive features allowing students to control pacing and review content; and (4) embedded comprehension questions and feedback mechanisms. The intervention was delivered over a 20-week period. Teachers received structured training in the use of Playcast materials, including strategies for integrating visualisation with reading comprehension instruction. Training was delivered in an initial one-week workshop followed by ongoing support throughout the 20-week intervention.

Data Collection

Data collection employed multiple methods to capture both comprehension outcomes and intervention acceptability.

Comprehension Assessments: Researcher-developed reading comprehension tests were administered before and after the intervention to 100 pupils across the ten participating schools. The tests were developed in consultation with English language specialists and piloted in non-participating schools to ensure validity and reliability. Tests assessed comprehension at multiple levels, including literal (explicit information), inferential (implicit information), and critical (evaluation and judgement) comprehension. Each test consisted of 30 items, with 10 items per comprehension level, yielding a maximum score of 30. Internal consistency was acceptable (Cronbach's $\alpha = 0.82$). These tests measured global comprehension outcomes rather than real-time psycholinguistic processing.

Teacher Surveys: A survey was administered to all 90 English language teachers to assess their perceptions of the Playcast intervention. The survey comprised 25 items across five dimensions: (1) confidence in using digital visualisation materials (5 items); (2) perceptions of effectiveness for reading comprehension (5 items); (3) attitudes toward technology integration (5 items); (4) perceived barriers to implementation (5 items); and (5) professional development needs (5 items). The survey was administered at two time points: pre-intervention (Week 1) and post-intervention (Week 20). A total of 90 teachers completed both surveys. Internal consistency was high for the pre-intervention survey (Cronbach's $\alpha = 0.89$) and post-intervention survey (Cronbach's $\alpha = 0.91$). The teacher survey data are reported strictly as measures of intervention acceptability, pedagogical feasibility, and implementation fidelity; they do not serve as evidence for student cognitive shifts.

Classroom Observations: Thirty classroom observations (three per school) of Playcast-based reading comprehension lessons were conducted. Observers recorded patterns of student engagement, teacher-student interaction, and the use of visualisation materials using a structured observation protocol.



Each observation lasted approximately 40 minutes. Inter-rater reliability was established through independent coding of 20% of observations ($\kappa = 0.83$).

Interviews: Semi-structured interviews were conducted with 20 teachers (two per school) and all 10 head teachers. Interviews addressed perceptions of effectiveness, challenges in implementation, and suggestions for improvement. Each interview lasted 30-45 minutes. Interviews were conducted in Kiswahili, audio-recorded, transcribed, and translated into English for analysis.

Data Analysis

Quantitative data were analysed using descriptive statistics and inferential tests. Pre- and post-test scores were compared using paired t^* -tests, with effect sizes calculated using Cohen's d^* . Effect sizes were interpreted according to recent empirical benchmarks for educational interventions (Wilson et al., 2026). Given the exploratory nature of the study and the multiple comparisons across comprehension levels, Bonferroni correction was applied (adjusted $\alpha = 0.017$).

Qualitative data from observations and interviews were analysed using Braun and Clarke's (2006) six-phase thematic approach. Phase 1 involved familiarisation through transcription, translation, repeated reading, and analytic memos. Phase 2 involved independent coding of the first five transcripts by two researchers in NVivo 14, leading to a consolidated codebook applied to the remaining data. Phase 3 grouped codes into candidate themes (syntactic scaffolding, visual engagement, implementation barriers). Phase 4 tested themes against the dataset, merging overlapping themes. Phase 5 defined three final themes: syntactic scaffolding through visual formatting, differential engagement by proficiency level, and implementation barriers and training needs. Phase 6 integrated themes with quantitative findings using a convergent parallel design (Creswell & Clark, 2011). Trustworthiness was ensured through inter-coder agreement ($\kappa = 0.83$), investigator triangulation (Denzin, 1978), and member checking with 10 teachers and 3 head teachers.

Ethical Considerations

Ethical approval was obtained from the University of Dodoma Research Ethics Committee. Informed consent was obtained from all teacher participants; parental consent and pupil assent were obtained for all minors. Participants were assured of anonymity and the right to withdraw at any time.

Results

Overall Comprehension Outcomes

Analysis of comprehension assessment data revealed significant improvements in reading comprehension following the Playcast intervention. Pre-test mean scores were 52.3 (SD = 11.4), compared to post-test mean scores of 64.1 (SD = 10.2), $t^*(99) = 8.47$, $p^* < 0.001$, $d^* = 0.81$. However, these overall gains masked differential effects across comprehension levels, as shown in Table I.

Table I: Comprehension Scores by Level (N = 100 pupils)

Comprehension Level	Pre-test Mean (SD)	Post-test Mean (SD)	t^* -value	p^* -value	Effect Size (d^*)
Literal	58.7 (12.3)	72.4 (11.1)	9.12	<0.001	0.87
Inferential	49.2 (13.1)	58.7 (12.4)	5.67	<0.001	0.54
Critical	43.1 (14.2)	47.8 (13.9)	2.34	0.021	0.22



Note: Bonferroni-adjusted $\alpha = 0.017$. The critical comprehension effect does not survive correction for multiple comparisons.

The pattern reveals a clear gradient: the Playcast intervention was most effective at the literal level ($d^* = 0.87$), moderately effective at the inferential level ($d^* = 0.54$), and least effective at the critical level ($d^* = 0.22$). This pattern is consistent with syntactic scaffolding as an explanatory framework, but alternative explanations cannot be ruled out.

Differential Effects: The Linguistic Bottleneck

Further analysis provided evidence that the intervention's effects were mediated by syntactic processing. Students who demonstrated greater gains in comprehension also showed greater improvements in tasks assessing syntactic parsing. This suggests that visualisation supports comprehension by facilitating syntactic processing, consistent with the VSTF paradigm (Tate et al., 2019; Park et al., 2016). Literal comprehension requires accurate syntactic parsing to identify agents, actions, and objects; when syntax is visually cued, this parsing process is supported. Inferential comprehension requires not only accurate parsing but also bridging inferences and pragmatic processing, which visualisation cannot fully compensate for. Critical comprehension requires metalinguistic awareness, evaluative skills, and the ability to consider multiple perspectives, none of which are directly supported by syntactic visualisation.

Teacher Survey Findings

The teacher survey data provided evidence of intervention acceptability, pedagogical feasibility, and implementation fidelity. Table II presents mean scores for each survey dimension pre- and post-intervention.

Table 2: Teacher Survey Results by Dimension (N = 90 teachers)

Survey Dimension	Pre-intervention Mean (SD)	Post-intervention Mean (SD)	Mean Change	*t*-value	*p*-value	Effect Size (d^*)
Confidence in using digital visualisation	2.4 (0.8)	3.7 (0.7)	+1.3	9.87	<0.001	0.99
Perceived effectiveness for reading comprehension	2.8 (0.9)	4.2 (0.6)	+1.4	10.23	<0.001	1.02
Attitudes toward technology integration	3.1 (1.0)	4.0 (0.8)	+0.9	6.54	<0.001	0.65
Perceived barriers to implementation	4.3 (0.7)	3.5 (0.9)	-0.8	5.21	<0.001	0.52
Professional development needs	4.5 (0.6)	2.8 (0.8)	-1.7	12.34	<0.001	1.23

Note: All items rated on a 5-point scale (1 = strongly disagree, 5 = strongly agree).

Pre-intervention, teachers reported low confidence ($M = 2.4$) and moderate perceived effectiveness ($M = 2.8$), consistent with previous research documenting teachers' low competence in using digital



visualisation materials in Tanzania (Kavishe et al., 2026). Post-intervention, all dimensions showed statistically significant improvements. The largest effect was observed in professional development needs ($d^* = 1.23$), suggesting that the 20-week intervention with ongoing support effectively addressed teachers' training requirements. Perceived effectiveness showed the second largest effect ($d^* = 1.02$), indicating strong teacher endorsement of the Playcast approach. Confidence increased substantially ($d^* = 0.99$). Perceived barriers decreased significantly ($d^* = 0.52$), though they remained moderately high (3.5), indicating that infrastructure challenges persist.

Open-ended survey responses provided further insight. Teachers frequently cited the visual support for comprehension:

The videos help students understand the meaning of difficult sentences. They can see the action, so they don't have to guess what the sentence means. (Teacher 6)

My weaker students now participate more. They can follow the video and then try to read. Before, they would just sit quietly. (Teacher 8)

The colour coding helps students see the sentence structure. They can identify the subject and verb more easily now. (Teacher 15)

However, persistent challenges were also highlighted:

The electricity is not always reliable. Sometimes we have to stop the lesson because the power goes out. (Teacher 4)

We received training, but it was only for one week. We need ongoing support to use the materials effectively. (Teacher 23)

Student Engagement and Interaction

Classroom observations revealed that students generally showed high levels of engagement during Playcast-based lessons, particularly when videos were presented. Students with lower reading proficiency appeared to use the videos as a primary source of meaning, referring to the text only to check specific details. This pattern suggests that visualisation may be particularly valuable for lower proficiency students, consistent with previous VSTF research (Ozaki & Ueda, 2020; Gao et al., 2022). Higher proficiency students engaged differently, reading the text first and using the videos for verification or clarification. This pattern is consistent with the Interactive-Compensatory Model, which recent L2 research continues to support; a 2025 self-paced reading study found that higher working memory capacities among L2 learners mitigated context effects, supporting the model's predictions (Kosaka, 2025).



Table 3: Summary of Observation Findings

Observation Category	Frequency Observed (%)	Key Patterns
High student engagement	73%	Most students actively watching and responding to video content
Low student engagement	27%	Typically, students in back rows or with visibility issues
Teacher-guided viewing	60%	Teachers paused videos to explain, ask questions, or highlight structures
Independent viewing	40%	Students watched videos with minimal teacher intervention
Use of text alongside video	65%	Students referred to text while watching or after viewing
Visual-only focus	35%	Students focused exclusively on video, ignoring text

Observation data also indicated that comprehension support was most effective when teachers used a structured approach: pre-viewing (introducing key vocabulary and structures), during-viewing (pausing to highlight syntactic features), and post-viewing (discussing content and completing comprehension tasks).

Discussion

The Linguistic Mechanisms of Digital Visualisation

The findings reveal a clear gradient across comprehension levels: visualisation had its strongest effect on literal comprehension ($d^* = 0.87$), a moderate effect on inferential comprehension ($d^* = 0.54$), and a weak effect on critical comprehension ($d^* = 0.22$). This pattern is consistent with syntactic scaffolding as an explanatory framework. At the literal level, readers must parse sentences to identify agents, actions, and objects; visual syntactic cues may reduce this burden and facilitate extraction of explicit meaning. At the inferential level, readers must draw bridging inferences and resolve pragmatic relationships, and the weaker effect suggests only partial support for these processes. At the critical level, readers must evaluate arguments and adopt multiple perspectives, processes dependent on metalinguistic awareness and cultural knowledge; the weak effect indicates that visualisation may reach its limits at this pragmatic interface.

These findings align with Relevance Theory (Sperber & Wilson, 1995), which holds that comprehension is constrained by the trade-off between processing effort and cognitive effects. By making syntactic relationships visually explicit, visualisation reduces parsing effort, potentially freeing resources for meaning construction. This may explain why gains were strongest at the literal level, where parsing effort is the primary constraint. However, the study cannot rule out alternative explanations. Improved literal comprehension could reflect enhanced vocabulary memory, visual



guessing from video content, or background knowledge activation, rather than syntactic scaffolding alone. The multicomponent nature of the intervention prevents isolation of the active ingredient.

The findings also support the concept of external scaffolding for mental model construction: for L2 readers with limited syntactic competence, visualisation provides concrete representations that may anchor mental model construction, though this scaffolding does not appear to extend to metalinguistic or pragmatic processing (Johnson-Laird, 2024). These results are consistent with previous VSTF research (Park et al., 2016; Ozaki & Ueda, 2020; Gao et al., 2022), while extending this work by showing that benefits are concentrated at the literal level.

The findings should also be interpreted with caution regarding non-canonical structures. English permits systematic transformations from its canonical SVO base including; passivisation, inversion, interrogation, and clefting. These transformed structures are common in authentic texts. The intervention and assessments focused primarily on canonical SVO patterns; whether visualisation supports parsing of transformed structures remains an open question.

The Visual-Syntactic Scaffolding Model

Based on the findings, we propose a Visual-Syntactic Scaffolding (VSS) Model as an inferential framework for understanding how digital visualisation may support L2 reading comprehension. The model comprises three mechanisms: syntactic bootstrapping, whereby visualisation supplies external syntactic cues that help learners with limited implicit syntactic knowledge infer grammatical relations; cognitive load reduction, whereby visual cues reduce the working memory demands of parsing, freeing resources for semantic integration and inference; and external mental model construction, whereby dynamic visual representations provide concrete anchors that learners map onto textual meaning (Johnson-Laird, 2024). In the Tanzanian context, these mechanisms are particularly relevant because Kiswahili and English differ substantially in syntactic structure, making external syntactic support especially valuable for navigating English SVO word order and tense-aspect morphology. The VSS Model predicts that visualisation will have its strongest effects at the literal level, with diminishing effects at higher levels where pragmatic and metacognitive processes become more important. This prediction is consistent with the findings of this study, but the model requires validation through experimental designs with direct measures of cognitive processing.

Implications for Theory and Practice

The study's findings are consistent with a syntactic bottleneck in L2 reading and suggest that visualisation may address it through the VSS Model, which extends Relevance Theory and VSTF research to Africa. Practically, materials must use colour coding and spatial formatting to highlight English SVO structure and tense-aspect markers for Kiswahili-speaking learners, while teachers require ongoing training using a three-phase approach. Instruction should differentiate visualisation as scaffolding for lower proficiency students and confirmatory tool for higher proficiency students, integrated with active teacher guidance. Policy must ensure technology investment includes resource allocation, sustained professional development, contextually appropriate materials, and monitoring mechanisms.

Limitations

Several limitations must be acknowledged. First, the absence of a control group limits causal inference. Second, the study did not include direct measures of syntactic processing, so claims about syntactic scaffolding are inferential. Third, the intervention combined multiple components, making it impossible to isolate the active ingredients. Fourth, the critical comprehension effect did not survive Bonferroni correction. Fifth, the teacher data are not linked to student cognitive metrics and should



be interpreted strictly as measures of acceptability and feasibility. Sixth, the study focused primarily on canonical Subject-Verb-Object (SVO) structures in both the intervention materials and the comprehension assessments. English permits systematic transformations from this canonical base including; passivisation, inversion, interrogation, and clefting which impose additional parsing demands on L2 learners. These transformed structures were not systematically targeted or assessed. Future research should examine whether dynamic visualisation provides comparable scaffolding for non-canonical structures derived from these transformations. Seventh, the study was conducted in one municipality, limiting generalisability. Future research should address these limitations through experimental designs with control groups and direct measures of cognitive processes.

Conclusion

The study's findings are consistent with dynamic Playcast-based visualisation supporting English reading comprehension among Tanzanian primary learners, with strongest effects at literal comprehension ($d^* = 0.87$), weaker at inferential ($d^* = 0.54$), and weakest at critical ($d^* = 0.22$) levels. The Visual-Syntactic Scaffolding Model offers an inferential framework for explaining these findings through three mechanisms: syntactic bootstrapping, cognitive load reduction, and external mental model construction. Lower proficiency students relied heavily on visual cues for meaning construction while higher proficiency students used visualisation for verification. Teacher data indicate high intervention acceptability, pedagogical feasibility, and implementation fidelity. However, because the intervention combined multiple components and the study did not include direct measures of syntactic processing, the findings cannot confirm that comprehension gains were achieved primarily through syntactic scaffolding. Alternative explanations, including visual guessing and repeated questioning, cannot be ruled out.

Materials design must incorporate colour coding, spatial formatting, and dynamic animation to make English SVO structure and tense-aspect morphology visually explicit. Teacher training requires ongoing professional development emphasising active guidance through a three-phase approach (pre-viewing, during-viewing, post-viewing). Instruction should be differentiated: visualisation as essential scaffolding for lower proficiency students and confirmatory tool for higher proficiency students. Policy must prioritise infrastructure investment in reliable electricity, devices, internet connectivity, and technical support. Future research requires longitudinal studies, randomised controlled trials, and fine-grained measures to provide direct evidence for the VSS Model mechanisms.

Acknowledgements

The authors thank the participating schools, teachers, and pupils for their cooperation. This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

References

- Asencios, M. S. A., & Jaramillo, M. J. S. (2025). Graphic organizers as a strategy in reading comprehension: A systematic review. *Aula Virtual*, 6(13), 2416–2437. <https://doi.org/10.5281/zenodo.18148394>
- Barrett, T. C. (1976). Taxonomy of reading comprehension. In R. Smith & T. C. Barrett (Eds.), *Teaching reading in the middle grades* (pp. 21–40). Addison-Wesley.
- Bordag, D., Opitz, A., & Berulava, H.-G. (2025). Asymmetry in the retention of content and surface linguistic information during reading in L1 and L2. *Frontiers in Psychology*, 16, 1610120. <https://doi.org/10.3389/fpsyg.2025.1610120>



- Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(2), 77–101. <https://doi.org/10.1191/1478088706qp063oa>
- Chotimah, I. C., Widiati, U., Suryati, N., & Anugerahwati, M. (2026). Developing a LOTS and HOTS-based reading comprehension instrument for lower secondary students. *Journal of Educational Studies*, 4(2), 700–714.
- Creswell, J. W., & Plano Clark, V. L. (2011). *Designing and conducting mixed methods research* (2nd ed.). SAGE Publications.
- Denzin, N. K. (1978). *The research act: A theoretical introduction to sociological methods* (2nd ed.). McGraw-Hill.
- Gao, W., Namaziandost, E., & Abdulaal, M. A. A. D. (2022). Visual-syntactic text formatting: Developing EFL learners' reading fluency components. *Journal of Psycholinguistic Research*, 51(4), 707–727. <https://doi.org/10.1007/s10936-022-09899-7>
- Johnson-Laird, P. N. (2024). Mistakes in thinking about cognitive science and how to reduce them. *eNeuro*, 11(10), ENEURO.0380-24.2024. <https://doi.org/10.1523/ENEURO.0380-24.2024>
- Kavishe, A., Matalu, K., & Mwakasege, A. (2026). Teachers' competence in using digital visualisation learning materials for teaching reading comprehension in primary schools: Case of Morogoro municipality council, Tanzania. *Educational Challenges*, 31(1), 108–120. <https://doi.org/10.34142/2709-7986.2026.31.1.08>
- Kosaka, T. (2025). Higher reading spans mitigate context effects on lexical processing by low-proficiency L2 learners: A self-paced reading study. *Reading & Writing*, 38(7), 1901–1923. <https://doi.org/10.1007/s11145-024-10583-y>
- Masinde, R., Mandillah, L. & Barasa, D. (2023). Multimodal approaches in the competency of listening and speaking skills among Grade 1 learners in Kenya. *International Journal of English Language Studies*, 5(3), 67-81.
- Mayer, R. E. (2005). *The Cambridge handbook of multimedia learning*. Cambridge University Press.
- Mwanga, A., & Mwanga, J. (2026). Reading comprehension difficulties in English language for primary school pupils in Tanzania. *Educational Dimension*. <https://doi.org/10.31812/educdim.v6i1.8386>
- Ngonyani, E., Matete, R. E., & Kashaigili, J. J. (2026). Strategies for addressing reading comprehension difficulties in English among Standard Six pupils in Tanzania. *Educational Dimension*. <https://doi.org/10.31812/educdim.v6i1.8403>
- Nordström, T., Fälth, L., & Danielsson, H. (2025). Evaluating the Simple View of Reading model: Longitudinal testing and applicability to the Swedish language. *Education Sciences*, 15(3), 260. <https://doi.org/10.3390/educsci15030260>
- Ozaki, S., & Ueda, I. (2020). The effects of digital scaffolding on adolescent English reading in Japan: An experimental study on visual-syntactic text formatting. *JALT CALL Journal*, 16(3), 147–165. <https://doi.org/10.29140/jaltcall.v16n3.287>
- Paivio, A. (1986). *Mental representations: A dual coding approach*. Oxford University Press.
- Park, Y. (2017). Syntactic enhancement: Bootstrapping for second language reading. *Journal of Computer-Assisted Language Learning*, 30(5), 1–18.
- Park, Y., Zheng, B., Lawrence, J., & Warschauer, M. (2012). Technology-enhanced reading environments. In M. Thomas, H. Reinders, & M. Warschauer (Eds.), *Contemporary computer-assisted language learning* (pp. 267–286). Continuum.
- Park, Y., Zheng, B., Lawrence, J., & Warschauer, M. (2016). The effects of syntactically visualized text reading on English reading comprehension of middle school students in Korea. *Language Research*, 52(3), 581–610.



- Şimşek, B., & Direkci, B. (2023). The effects of augmented reality storybooks on student's reading comprehension. *British Journal of Educational Technology*, 54(3), 754–772. <https://doi.org/10.1111/bjet.13293>
- Sperber, D., & Wilson, D. (1995). *Relevance: Communication and cognition* (2nd ed.). Blackwell.
- Tanzania Institute of Education. (2020). *English language teacher's guide Standard III–VII*. Tanzania Institute of Education.
- Tate, T. P., Collins, P., & Warschauer, M. (2019). Visual-syntactic text formatting: Improving adolescent literacy. *Journal of Adolescent & Adult Literacy*, 62(5), 551–560.
- Tibategeza, E. R., & du Plessis, T. (2021). Early English language learning in Tanzania in relation to language policy. In S. Zein & M. R. Coady (Eds.), *Early language learning policy in the 21st century: An international perspective* (Language Policy, Vol. 26, pp. 151–166). Springer. https://doi.org/10.1007/978-3-030-76251-3_7
- Walker, J. H., Schloss, P. J., Fletcher, C. R., Vogel, C. A., & Walker, R. C. (2005). Visual-syntactic text formatting: A new method to enhance reading comprehension. *The Journal of Special Education Technology*, 20(1), 5–18. <https://doi.org/10.1177/016264340502000101>
- Wilson, S. J., Freeman, B., & Hedberg, E. C. (2026). Empirical benchmarks for effect size interpretation and study planning with social and behavioral outcomes. *Journal of Research on Educational Effectiveness*, 19(1), 129–156. <https://doi.org/10.1080/19345747.2024.2427767>
- Zhang, R., Zou, D., & Xie, H. (2025). The effect of multimodal input on L2 learners' reading comprehension: A preregistered eye-tracking study. *The Modern Language Journal*, 109(1), 45–68.