

MULTIMODAL EXPERIENTIAL MEANING IN NINTH-GRADE DIGITAL PROCEDURE INFOGRAPHICS: A SYSTEMIC FUNCTIONAL LINGUISTIC ANALYSIS

¹*Lina Ekawati, ¹Dwi Santoso, ¹Surono

¹English Language Education, Faculty of Teacher Training and Education, Universitas Ahmad Dahlan, Daerah Istimewa Yogyakarta Province, Indonesia

*Corresponding Author Email: 2575042011@webmail.uad.ac.id

Article Info	Abstract
Article History Received: June 2026 Revised: August 2026 Accepted: September 2026 Published: October 2026	<i>While digital compositions are increasingly prominent in language education, the unresolved problem remains of how student-created digital infographics coordinate language, image, and layout to represent procedural experience. To address this gap, this qualitative case study investigates the multimodal realization of experiential meaning in digital procedure infographics created by ninth-grade students in a private junior high school within a Central Kalimantan palm oil plantation, Indonesia. Analyzing ten purposive student infographics using Systemic Functional Linguistics (SFL) Transitivity and Visual Grammar, data were triangulated with semi-structured interviews and observations. Findings reveal that verbally, instructional steps are heavily dominated by Material Processes (90%–95%), transitioning to Relational Processes at closures to signal product readiness. Visually, students localized the genre by integrating Dayak cultural motifs to assert local identity. Intermodally, a critical "Static Paradox" (intermodal mismatch) occurs when dynamic verbal commands are paired with conceptually frozen static images lacking action vectors, a challenge further constrained by unstable plantation Wi-Fi. This study contributes directly to English language teaching by demonstrating how socio-technical and cultural factors shape students' multimodal representations. To guide layout, text-image alignment, and culturally situated designs, educators should implement an Integrated Multimodal Cohesion Rubric for assessment and introduce Structural Macro-Layout Blueprints during the design process.</i>
Keywords Multimodal analysis; Experiential meaning; Infographics; Procedure texts; Systemic functional linguistics;	

How to cite: Ekawati, L., Santoso, D., & Surono, D.S. (2026). Multimodal Experiential Meaning in Ninth-Grade Digital Procedure Infographics: A Systemic Functional Linguistic Analysis. *JOLLT Journal of Languages and Language Teaching*, 14(4), 1299-1314. Doi: <https://doi.org/10.33394/joltt.v14i4.21374>

Copyright© 2026, Lina Ekawati et al.

This is an open-access article under the [CC-BY-SA](#) License.



INTRODUCTION

Language has long been recognized as a fundamental social semiotic system; however, in the rapidly evolving contemporary educational landscape, the process of meaning-making has become increasingly multimodal (Bezemer & Kress, 2016; Pauwels & Mannay, 2019). This pedagogical evolution is marked by a crucial shift from traditional, purely linguistic systems toward complex interactions among various semiotic resources, including images, video, text, and sound, especially within the realm of digital student work. Such multimodal artifacts serve as a powerful vessel for constructing and expressing experiential meaning in the modern classroom. In this framework, experiential meaning is realized through the strategic ways in which students

choose and compose different modes to represent and organize their unique understanding of the world. This realization involves the detailed portrayal of participants, processes, and circumstances, which collectively constitute the representation of human experience (Huang et al., 2021). Within the Indonesian context of Kurikulum Merdeka, there is an increasing emphasis on digital literacy and the ability to communicate effectively through multiple semiotic resources, prompting English Language Teaching (ELT) practitioners to transition from traditional essays to multimodal production tasks (Sadhu & Ghasani, 2023). Among various digital formats, infographics have emerged as a prevalent and valuable medium in ELT, particularly for presenting procedural texts, which require students to sequence actions and information logically and clearly (Jewitt, 2017).

Despite the theoretical benefits of multimodal compositions, a significant empirical and pedagogical gap remains regarding how students actually navigate the complex integration of visual and textual resources to achieve instructional coherence. Existing literature predominantly focuses on professional or published materials, such as textbook dialogues (Sugianto et al., 2022) or celebrity Instagram posts (Indrayani, 2018), rather than authentic student-generated artifacts. Furthermore, previous studies often prioritize traditional linguistic accuracy (Mulki et al., 2022) or analyze modes in isolation within spoken discourse (Megah, 2019; Kyaw, 2024), leaving a lack of empirical evidence regarding how language learners simultaneously deploy verbal and visual transitivity resources to construe procedural experiences (Putri, 2024).

This study addresses these gaps by defining the context of situation and culture through the unique setting of SMP Eka Tjipta Semilar, a private junior high school situated within an isolated palm oil plantation in Central Kalimantan, Indonesia. This remote environment acts as a crucial variable that directly shapes and constrains students' multimodal choices: their everyday plantation experiences influence their procedural topics and ingredient selections, their limited digital access and unstable Wi-Fi constrain their visual layout options, and their integrated visual elements such as Dayak motifs, regional wildlife (e.g., hornbills and orangutans), and traditional houses represent a deliberate semiotic strategy to express and negotiate local cultural identity on the digital canvas. Grounded in the framework of Systemic Functional Linguistics (SFL) developed by Halliday (2014) specifically Halliday and Martin's concept of the Context of Culture, this study argues that cultural ornamentation in learner-generated texts is not merely decorative, but functions as a critical interpersonal vehicle for cultural positioning (Martin, 2024).

To systematically decode these digital artifacts, this study integrates SFL Transitivity theory with Visual Grammar. In SFL, the experiential metafunction models real-world experiences through clauses configured as processes, participants, and circumstances (Halliday & Matthiessen, 2014). Verbally, this is realized through transitivity processes (primarily Material, Relational, and Mental) that categorize our experiences of the world. Visually, this framework is extended through Kress and van Leeuwen's Visual Grammar (2021), which categorizes visual processes into narrative structures (depicted through explicit or implied Action Vectors representing movement) and conceptual structures (depicted as static states of classification or identification), alongside visual participants and visual circumstances of setting.

The novelty of this study lies in its holistic integration of authentic student-generated compositions, verbal-visual transitivity analysis, local cultural identity representation, and rural technological conditions within a single qualitative case study. While earlier research has investigated these elements in isolation or focused on resource-rich urban classrooms, this research demonstrates how marginalized EFL students navigate the tension between standardized digital curriculum demands and remote, localized realities. By investigating how environmental

limitations (e.g., slow plantation Wi-Fi) and cultural motives interact with visual-verbal texturing, this study provides a fresh perspective on authentic digital literacy that moves beyond traditional text-centric, error-focused grading

In alignment with these theoretical and contextual orientations, this research frames its investigation around two primary questions, specifically exploring how experiential meaning is realized multimodally through the interplay of text, images, and layout in the ninth-grade students' digital procedure infographics at a private junior high school in Central Kalimantan, as well as identifying the prominent patterns, strategies, and challenges observed in the students' realization of transitivity components in their digital procedure infographics.

RESEARCH METHOD

Research Design

This research methodology is structured as a qualitative case study designed to provide a "thick description" of multimodal writing and digital literacy practices within a localized educational setting. To maintain methodological rigor, this case study is strictly bounded by institutional, classroom, temporal, and task parameters. Institutionally, the study was conducted at SMP Eka Tjipta Semilar, a private junior high school located within an isolated palm oil plantation in Central Kalimantan, Indonesia, which operates under the Eka Tjipta Foundation (Sinar Mas Group). Classrooms were bounded to Ninth-Grade Class IX-1 (30 students) and Class IX-2 (30 students). The digital task under investigation was the individual creation of a procedural text infographic on the Canva design platform as a formative assessment. Temporally, the research spanned the second semester of the 2025/2026 academic year, with the field data collection phase conducted from February to April 2026.

Chronologically, the infographic project was embedded within a structured 9-hour (9 JP, where 1 JP = 40 minutes) instructional sequence spanning three weekly meetings in April 2026 based on the official lesson plans (Modul Ajar). Meeting 1 (3 JP) introduced procedural structures, transitivity components, and multimodal models using video and text. Meeting 2 (3 JP) engaged students in worksheet activities (LKPD) and initiated individual digital drafting on Canva. Meeting 3 (3 JP) concluded with digital submissions to the virtual learning portal (belajar.id), oral summative assessments, and the execution of retrospective interviews.

Research Participants and Sampling Strategy

The participants comprised sixty ninth-grade EFL students aged fourteen to fifteen, split equally across Class IX-1 and Class IX-2 at SMP Eka Tjipta Semilar, Central Kalimantan, situated within an isolated palm oil plantation. Utilizing a purposive sampling strategy to preserve analytical depth, ten infographics (five from each class) were selected based on predetermined criteria. The inclusion criteria required complete procedural structures, at least three steps, and corresponding visual elements per step; conversely, the exclusion criteria disqualified incomplete, non-English, or purely text-based decorative designs. Selected artifacts meeting these criteria were randomized using a number generator. To triangulate findings, retrospective semi-structured interviews were conducted with the ten selected student creators (establishing a direct analytical relationship) and their English subject teacher.

Research Objects and Instruments

The primary object of this study is the realization of experiential meaning (transitivity) through the interplay of linguistic and visual elements in student-generated digital procedure infographics. In accordance with qualitative case study standards, the primary research instrument is the researcher, who exercises interpretive rigor to bridge abstract SFL theory with

visual and verbal data. Supporting instruments include the Multimodal Analytical Matrix (used as the coding guide to analyze verbal transitivity, visual action vectors, and inter-modal relations) and the semi-structured Interview Guide containing targeted protocols for students and the teacher. Other components, such as the digital files of student infographics, classroom observation field notes, and the data anonymization protocol, are classified as qualitative data records and procedures rather than research instruments.

Data Collection Techniques

Data were collected through a triangulated, multi-tiered qualitative process to capture both the final artifacts and the contextual classroom environment. First, the primary data source was the ten high-quality digital files of student-generated Canva infographics, gathered electronically to preserve complete visual integrity (including colors, typography, and spatial resolutions). Second, classroom observations were recorded in a reflective journal and field notes during Session 2 to document the physical plantation infrastructure, available digital tools, and the direct pedagogical scaffolding provided by the teacher. Third, retrospective, semi-structured interviews were recorded and transcribed, asking the ten student creators about their choices of visual salience, color, layout, and image-text alignment, and asking the English teacher about their pedagogical goals and grading priorities.

Data Verification and Trustworthiness

To establish the trustworthiness, credibility, and dependability of the qualitative findings, three main verification techniques were implemented. First, Method Triangulation was rigorously executed by systematically cross-checking and comparing data from the student infographics (document analysis), classroom observations (reflective journal), and participant interview transcripts to check for interpretive convergence or divergence. Second, Thick Description was provided by detailing the exact analytical frameworks, coding categories, and sociocultural context of the palm oil plantation, backed by anonymized direct quotes from interviews and concrete visual examples from student work. Third, Expert Judgment was utilized to eliminate researcher bias; the detailed coding sheets, intermodal matrices, and thematic interpretations were presented to and verified by the thesis supervisors, who acted as external experts in Multimodal Discourse Analysis (MDA) and Systemic Functional Linguistics (SFL).

Data Analysis Procedures

The data analysis in this study follows a Multimodal Discourse Analysis (MDA) approach utilizing Miles, Huberman, and Saldaña's (2014) interactive model of data condensation, data display, and conclusion drawing. Systematically, the tripartite analytical coding was executed recursively across all ten selected infographics. First, verbal texts were segmented into clauses and coded using Systemic Functional Linguistics (SFL) transitivity categories to identify processes, participants, and circumstances. Second, visual elements were coded using Kress and van Leeuwen's (2021) Visual Grammar, distinguishing between dynamic Narrative Processes (action vectors) and static Conceptual Processes. Third, intermodal relations were mapped using a Multimodal Analytical Matrix to determine logico-semantic complementarity. Coded features were then compared in a cross-artifact matrix to synthesize recurring patterns, such as the Relational Shift and the Static Paradox, across the ten infographics. To ensure reliability, the primary researcher acted as the first analyst, while a peer analyst independently double-coded thirty percent of the corpus. Coding discrepancies were resolved through collaborative recoding sessions until reaching a hundred percent consensus, which was subsequently verified by the thesis supervisors acting as external experts in SFL and MDA.

RESEARCH FINDINGS AND DISCUSSION

Research Findings

This section presents the results of the multimodal analysis conducted on ten digital procedure infographics created by ninth-grade English as a Foreign Language (EFL) students at SMP Eka Tjipta Semilar, located within a remote palm oil plantation in Central Kalimantan, Indonesia. The findings reveal that these secondary students construct experiential meaning through a complex, functional, and culturally situated orchestration of linguistic, visual, and spatial resources. To establish a systematic empirical baseline, this section begins with a corpus-wide synopsis of cross-artifact patterns across all ten analyzed infographics, which are summarized in Table 1.

The corpus-wide analysis reveals distinct patterns in the verbal, visual, and intermodal realization of transitivity components. Verbally, the instructional steps are heavily dominated by Material Processes, which constitute approximately 90% to 95% of all verbal clauses in the core steps. This absolute dominance represents the procedural experience as a chronological sequence of physical actions (e.g., mix, pour, stir, boil, sauté, wash). These processes consistently act upon participant roles functioning as Goals, which represent the ingredients, tools, or objects physically affected by the actions, while the Actor (the person performing the procedure) is left implied through the imperative mood. To achieve instructional precision, students regularly employ Circumstances of Location: Place, Manner: Quality, and Contingency: Condition to guide the reader. A unique linguistic pattern observed is the "Relational Shift" at the closure of the procedural cycle, where the final clause transitions to a Relational Process to attribute a state of completion or readiness to the finished product.

Visually, the infographics rely primarily on static Conceptual Structures to identify ingredients, materials, and finished products, representing participants in terms of their stable essence, class, or structure. However, high-performing samples successfully deploy Narrative Processes characterized by explicit visual Action Vectors, such as directional arrows and circular lines, to depict physical movement. A prominent visual strategy is the "Localization Strategy," where students systematically embed Dayak and Borneo cultural motifs including hornbills, orangutans, traditional houses, dancers, and ornate border patterns as visual participants. These motifs function as Conceptual Classificational Processes, anchoring the global digital genre in the students' immediate regional heritage and lived experience. Additionally, students employ a "Hybrid Modality" strategy, pairing high-modality realistic photographs for the final Goal with simplified, stylized icons for intermediate steps to balance visual appeal with spatial efficiency.

Intermodally, the dominant pattern is "Inter-semiotic Complementarity" realized through elaborative and additive relations. Visual elements complement the verbal sequence by providing immediate identification of materials and structural layout organization. However, a critical intermodal breakdown emerged as the "Static Paradox" (Intermodal Mismatch). This paradox occurs when dynamic, action-oriented verbal commands are paired with conceptually frozen static images lacking action vectors or directional lines of motion. Triangulated data from reflective journals and interviews confirms that this paradox is frequently triggered by technical and environmental constraints. Slow and unstable Wi-Fi connectivity in the remote palm oil plantation setting appeared to influence student choices, often forcing them to abandon complex, logically aligned visual designs and settle for basic, pre-packaged icons to complete the Canva task within the allotted time. Additionally, students often prioritized commercial-looking aesthetics over procedural navigation, leading to high visual density and decorative distractors

that obscured the reading path. Table 1 provides a consolidated synopsis of the entire ten-artifact corpus.

Table 1
Multimodal Realization of Experiential Meaning

Artifact Code	Procedural Topic	Verbal Process Distribution	Visual Process Type	Intermodal Relationship	Cultural Motifs	Observed Design Problems
9-1.ST-13	Chocolato Moss Ice	Material (100% of steps)	Conceptual	Elaborative, Additive	Dayak motifs, hornbills, orangutan, traditional house	Excessive visual noise from dense cultural graphics, high information density
9-1.ST-14	Chocolato Pudding	Material (100% of steps)	Conceptual	Elaborative, Additive	Dayak ornaments, dancers, hornbill bird	Busy background and complex layout distracting from instructional steps
9-1.ST-28	Es Pisang Ijo Borneo	Material (100% of steps)	Conceptual	Elaborative, Additive	Hornbill bird	"Conceptually frozen" static images, high text reliance, dense pink background
9-1.ST-20	Chicken Noodles	Material (100% of steps)	Conceptual	Elaborative, Additive	Ornate traditional Dayak/Borneo border style	Intermodal Mismatch (Static Paradox), lack of action vectors, dense visual header
9-1.ST-18	Meatballs with Beef Tendon	Material (100% of steps)	Conceptual	Elaborative, Additive	None (Culinary characters and chef icon)	Static imagery, cognitive hurdle translating "Sauté" to icons, missing ingredients
9-2.ST-07	Traditional Style Slime	Material (90%), Relational (10% closure)	Conceptual	Elaborative, Additive	Ornate blue patterns (local aesthetics)	Static imagery, lack of action for "knead" and "stir" steps
9-2.ST-01	Ice Kusuu (Grape Drink)	Material (90%), Relational (10% closure)	Conceptual	Elaborative, Additive, Congruent	Ornate purple layout accents (grape theme)	Spatial disorientation, asymmetrical layout forcing diagonal visual leaps, fractured path
9-2.ST-05	Dalgona Coffee	Material (90%), Relational (10% closure)	Conceptual	Elaborative, Additive	Dayak dancers, hornbills, traditional patterns	"Conceptually frozen" static images for action steps like "stir" or "whisk"
9-2.ST-30	Strawberry Ice	Material (90%), Relational (10% closure)	Conceptual	Elaborative, Additive	None (Commercial aesthetic, pink bow, pastel tones)	Intermodal Mismatch (Static Paradox), dynamic verb "Pour" paired

Artifact Code	Procedural Topic	Verbal Process Distribution	Visual Process Type	Intermodal Relationship	Cultural Motifs	Observed Design Problems
9-2.ST-24	6 Hand Washing Steps	Material (100% of steps)	Narrative (with explicit action vectors)	Elaborative, Additive, Congruent	Dayak cultural patterns, ornate borders, Dayak warrior	with a static finished cup photo Numbering reset (1-2-3 repeated twice), compensated by clear Z-pattern layout

To illustrate the broader patterns and tensions identified in the corpus-wide analysis, two contrasting case studies are presented in detail below. The first case study, Infographic 9-1.ST-13 (Chocolato Moss Ice), represents the dominant static conceptual recipe pattern, which is characterized by rich local cultural ornamentation and a high-modality realistic product photograph. The second case study, Infographic 9-2.ST-24 (6 Hand Washing Steps), represents the dynamic narrative pattern, which successfully employs explicit visual action vectors and structural layout framing to guide step-by-step instructional flow.

Case Study 1: The Infographic 9-1.ST-13 (Chocolato Moss Ice)

Infographic 9-1.ST-13 illustrates the broader pattern of recipe-focused digital student work that prioritizes local cultural aesthetics and product identification over narrative-procedural motion, creating a clear tension between linguistic dynamics and visual stasis. Linguistically, the infographic's steps are realized through Material Processes that outline the physical transformation of ingredients. Table 2 provides a systematic clause-by-clause SFL transitivity analysis of this artifact, separating the lexical verb from the SFL process category and using standard SFL terminology.

Table 2
Verbal Analysis of Infographic 9-1.ST-13

Generic Structure	Clause	Process Type	Participant	Circumstance
Goal	CHOCOLATO MOSS ICE	(Implied: make)	Chocolato Moss Ice	-
		Material	Goal	-
Materials	chocolate agar-agar, sugar, egg, water, etc.	-	-	-
		-	conceptual	-
Step 1	Mix the chocolate agar-agar powder, sugar, and water in a pot.	Mix	chocolate agar-agar powder, sugar, and water	in a pot
		material	goal	location
Step 2	Add 1 egg and stir well.	Add, stir	1 egg	well
		material	goal	manner
Step 3	Cook over medium heat while stirring continuously.	Cook, stirring	-	over medium heat, continuously
		material	-	Condition, manner
Step 4	...reduce the heat and stir slowly until moss-like strands form.	reduce, stir	the heat	Slowly, until moss-like strands form.
		material	goal	manner, condition
Step 5	...refrigerate until set.	refrigerate	-	until set
		material	-	condition
Step 6	Take some chocolate moss and put it into a clear cup.	Take, put	chocolate moss, it	into a clear cup
		material	goal	location

Source: Researcher's verbal analysis

The SFL analysis in Table 2 reveals that the verbal text operates as a highly cohesive series of physical actions. Common command verbs function as the core Material Processes acting upon participant roles functioning as Goals. The Actor remains implied through the imperative mood, aligning with the standard functional grammar of procedural recipes. To provide high technical precision, the student incorporates multiple Circumstances of Location: Place, Manner: Quality, Manner: Means, and Contingency: Condition. This linguistic structure successfully frames the procedural world as a functional, chronological sequence of physical doings. Visually, however, the infographic operates as a static conceptual structure rather than a dynamic narrative process. Table 3 lists the visual elements utilized in Infographic 9-1.ST-13 and describes their representational functions.

Table 3
Visual Analysis of Infographic 9-1.ST-13

Visual Element	Type	Representational Function
Photo of Moss Ice in Bottles	Graphic (Photo)	Conceptual participant (goal). Represents the desired finished product with high modality, showing the texture of the "moss".
Chocolate Bar	Graphic	Conceptual participant (material). Visually reinforces the key ingredient (chocolate) listed in the text.
Borneo/Dayak Imagery (Hornbills, Orangutan, Traditional House)	Graphic	Conceptual classification (identity). Anchors the text in a specific cultural and geographic context (Borneo/Central Kalimantan), giving the recipe a unique origin.
Scalloped Frames	Layout	Structural organization. Acts as visual dividers to separate the ingredients from the procedural steps, making the hierarchy clear.
Bold Serif Typography	Typography	Emphasis (hierarchy). Large fonts for "CHOCOLATO MOSS ICE" and "Ingredients" signal importance and guide the user through the segments.

Source: Researcher's visual analysis

As shown in Table 3, the visual mode is dominated by Conceptual Participants that represent identity and states of being. The student uses a high-modality realistic photograph of the finished drink at the top-left to establish the concrete Goal of the procedure. More importantly, the student deploys a strong "Localization Strategy" by integrating rich Borneo and Dayak cultural graphics. These illustrations function as Conceptual Classification structures that establish cultural identity. While these elements perform vital interpersonal work by asserting local heritage on a digital canvas, they create an extremely high visual density. This results in visual noise that appeared to increase processing demands and threatens to distract from the procedural steps, though the student compensates for this by placing the text blocks inside high-contrast white frames.

Intermodally, the relationship between text and visual modes in Infographic 9-1.ST-13 is characterized by elaborative and additive complementarity, but it also reveals a critical intermodal breakdown. Table 4 outlines these intermodal interactions.

Table 4
Intermodal Analysis of Infographic 9-1.ST-13

Multimodal Feature	Verbal	Visual	Intermodal Relationship
The Goal	The text "CHOCOLATO MOSS ICE"	It is paired with a realistic photograph of the drink.	Elaborative (additive): The image illustrates the abstract goal, showing the reader exactly what the finished "moss" texture should look

Multimodal Feature	Verbal	Visual	Intermodal Relationship
			like.
The materials	The “material”: chocolate agar-agar, sugar, egg, water, etc.	The ingredients are housed in a white bubble with a brown, scalloped outline.	Additive (identification): The verbal mention of chocolate is reinforced by a graphic of a chocolate bar at the bottom, identifying a key starting material.
Structure & Flow	The heading: Goal, Material, Steps	The use of distinct decorative frames for the ingredients and steps mirrors the verbal organization	Additive (structural): This ensures the reader experiences the process as a clear, sequential flow.
Overall theme	The repeated mention of ‘chocolate, milk, water’	The overall brown and canvas texture	Congruent association: While the coffee example used brown to evoke the drink itself, the writer uses rich cultural illustrations (Dayak patterns and local wildlife) to associate the recipe with a specific regional identity, making the experience feel authentic to her background.

Source: Source: Researcher’s intermodal analysis

Table 4 highlights the core intermodal synergy and tension of this pattern. While the color semiotics and layout framing support structural cohesion through Congruent Association and Additive Structural relationships, the actual procedural steps suffer from the "Static Paradox". For example, when the verbal text commands the reader to "stir well" or "cook while stirring continuously," the accompanying graphics are static and lack action vectors. The reader must mentally bridge the gap between the dynamic verbal action and the frozen visual state, indicating that visual resources in this recipe infographic are utilized for aesthetic and cultural classification rather than step-by-step physical demonstration.

Case Study 2: Infographic 9-2.ST-24 (6 Hand Washing Steps)

Infographic 9-2.ST-24 represents the contrasting dynamic narrative pattern, which successfully coordinates dynamic verbal commands with visual action vectors and strict layout framing, emphasizing visual-verbal action-oriented alignment. Linguistically, the verbal mode is composed entirely of Material Processes. Table 5 provides the systematic SFL transitivity analysis of this artifact, separating the lexical verb from the SFL process category and using standard SFL terminology.

Table 5
Verbal Analysis of Infographic 9-2.ST-24

Generic Structure	Clause	Process Type	Participant	Circumstance
Goal	6 Hand Washing Steps as recommended	-	-	-
Recommendation	Wash your hands properly...	Wash material	your hands goal	properly manner
Step 1	rinse hands with clean water.	rinse material	hands goal	with clean water manner
Step 2	use enough soap.	Use material	soap goal	enough extent
Step 3	rub your palms thoroughly.	rub material	your palms goal	thoroughly manner
Step 4	rub the backs of your hands...	rub	backs of your hands	-

Generic Structure	Clause	Process Type	Participant	Circumstance
		material	goal	-
Step 5	rub your thumb in a circular motion.	rub	your thumb	in a circular motion
		material	goal	manner
Step 6	rinse with water, then dry.	Rinse, dry	-	with water
		material	-	Manner

Source: Researcher's verbal analysis

As shown in Table 5, the verbal text is realized almost exclusively through Material Processes acting upon participant roles functioning as Goals. The Actor remains implied through the imperative mood. To provide instructional clarity, the student incorporates multiple Circumstances of Manner: Quality, Manner: Means, Extent: Amount, and Location: Time. This grammatical configuration frames the hygiene procedure as a highly functional, logical sequence of actions. Visually, this infographic successfully deploys Narrative Processes. Table 6 provides the visual grammar analysis of this artifact, outlining how the visual elements represent experiential meaning.

Table 6
Visual Analysis of Infographic 9-2.ST-24

Visual Element	Type	Representational Function
Tap, Sink, and Hands	Graphic	Conceptual participant (goal). A clear visual representation of the subject matter at the top, setting the scene for the procedure.
Instructional Action Boxes	Graphic	Conceptual participant (narrative). Six blue boxes showing the specific hand positions, acting as visual "Action Vectors" for the verbal steps.
Dayak Cultural Patterns	Graphic	Conceptual classification (identity). The ornate red, yellow, and blue borders and the Dayak warrior figure ground the global health advice in a local cultural context.
Large White "6"	Typography	Emphasis (hierarchy). Signals the number of steps to follow, guiding the user through the logical progression.
Blue Background & Blocks	Color/Layout	Structural organization. The blue color evokes a sense of cleanliness and water, while the blocks divide the sequence into digestible parts.

Source: Researcher's visual analysis

As described in Table 6, this infographic differs from the first case study by utilizing Narrative Processes. In the six instructional panels, the student depicts hand positions with explicit Action Vectors such as circular motion arrows for rubbing thumbs and back-and-forth lines for rubbing palms. These visual vectors directly mirror and reinforce the dynamic verbal commands, successfully overcoming the Static Paradox. Structurally, the blue background establishes a thematic association with water and cleanliness, while the segmented block layout divides the steps into digestible framed steps. The student also localizes the global health advice by incorporating ornate Dayak borders and a traditional warrior figure, grounding the digital artifact in local identity without cluttering the core step zones. Intermodally, the relationship in Infographic 9-2.ST-24 is characterized by highly functional, elaborative, and structural synergy. Table 7 outlines the intermodal interactions of this artifact.

Table 7
Intermodal Analysis of Infographic 9-2.ST-24

Multimodal Feature	Verbal	Visual	Intermodal Relationship
The Goal	The text "6 Hand Washing Steps"	It is paired with a large graphic of hands being washed under a tap	Elaborative (additive): The image illustrates the abstract concept of the goal, representing the final state or the

Multimodal Feature	Verbal	Visual	Intermodal Relationship
			general activity described verbally.
The Steps	Each verbal instruction (e.g., "rub your palms")	It is paired with a corresponding illustration of that action	Additive (structural): The visual layout directly mirrors the verbal structure, ensuring a clear sequential flow.
Overall Theme	The use of water-themed blue colors	reinforces the central theme of washing	Congruent association: The traditional Dayak imagery makes the experience feel authentic to the creator's regional identity.

Source: Researcher's intermodal analysis

As shown in Table 7, this infographic achieves a high level of inter-semiotic complementarity. Every written step is paired with an illustrated panel that actively clarifies and represents the verbal action visually. The visual mode does not merely repeat the text; it represents the doing of the steps through explicit motion vectors. However, a specific navigational challenge was observed in this design. Despite the title stating "6 steps," the illustrated panels are numbered as "1, 2, 3" in the top row and "1, 2, 3" again in the bottom row due to a numbering reset on Canva. This layout choice could potentially increase processing demands. Nevertheless, the student successfully compensates for this potential breakdown by deploying a clear Z-pattern layout, relying on the reader's natural scanning path (top-to-bottom and left-to-right) to maintain the chronological and logical flow of the procedure. This design balance illustrates that when students are provided with structured layout frameworks, they can effectively negotiate the tension between digital aesthetics and functional instructional logic.

Discussion

The findings demonstrate that ninth-grade EFL students construct experiential meaning in digital procedure infographics through an uneven but functionally significant orchestration of linguistic, visual, spatial, and cultural resources. The most prominent linguistic finding is the overwhelming dominance of Material Processes, accounting for approximately 90–95% of the clauses within the instructional steps. Verbs such as *mix*, *pour*, *stir*, *cook*, *wash*, and *rub* construe procedural experience as a sequence of concrete transformations performed upon particular Goals, while the Actor is generally omitted through the imperative structure. This pattern is consistent with Halliday and Matthiessen's (2014) account of transitivity, in which Material Processes construe physical events and actions in the experiential world. It is also compatible with the generic characteristics of procedural discourse described by Knapp and Watkins (2005), where imperative clauses and action verbs organize instructions chronologically. However, the present findings extend previous research on students' procedure texts, which has frequently concentrated on linguistic errors and grammatical accuracy (Mulki et al., 2022), by demonstrating that procedural competence also involves the multimodal realization of actions. Thus, successful procedural meaning is not determined exclusively by whether learners select grammatically appropriate action verbs, but also by whether these actions are coherently represented through other semiotic modes.

The visual findings, however, reveal an important asymmetry between linguistic and pictorial representations. Whereas the verbal mode is predominantly dynamic, most visual elements are realized through static Conceptual Structures representing ingredients, tools, cultural symbols, or finished products. From the perspective of Visual Grammar, conceptual representations identify what participants *are*, whereas narrative representations depict what participants *do* through vectors indicating action or movement (Kress & van Leeuwen, 2021). Consequently, the predominance of static images in these infographics indicates that many

students treated visual resources primarily as objects of identification or decoration rather than as resources for representing procedural action. This finding supports the broader argument that multimodal competence cannot be assumed simply because learners can use digital applications. Digital composing requires learners to understand the distinct affordances of different semiotic resources and coordinate them meaningfully (Bezemer & Kress, 2016; Cope & Kalantzis, 2020). O'Halloran (2008) similarly argues that multimodal meaning emerges not merely from the coexistence of language and visual imagery but from the relationships established across semiotic systems.

This tension becomes particularly visible in what the present study identifies as the Static Paradox. Dynamic commands such as *stir continuously*, *knead*, *pour*, or *rub* are frequently paired with static illustrations that show ingredients, containers, or completed products without vectors indicating the required movements. In these cases, verbal and visual resources coexist but do not fully cooperate in representing the same experiential event. The reader must infer the missing action and mentally reconstruct the relationship between the command and the image. This finding is broadly consistent with Liu and O'Halloran's (2019) argument that intersemiotic relations determine how meanings are expanded, specified, or reorganized across semiotic resources. It also qualifies the assumption that simply adding images makes instructional texts clearer. Images become pedagogically useful only when their representational function corresponds to the communicative demands of the genre. In a procedural infographic, the critical function of a visual is often not merely to identify an object but to demonstrate transformation, direction, sequence, or movement.

The contrasting case of *6 Hand Washing Steps* provides strong evidence for this interpretation. Unlike most recipe-based artifacts, this infographic employs explicit vectors, circular arrows, hand positions, segmented instructional panels, and a relatively stable Z-pattern reading path. As a result, the images represent the actual *doing* described verbally instead of functioning primarily as decorative illustrations. Each verbal command is paired with a visual realization of the relevant action, producing stronger elaborative and structural complementarity. Although the duplicated numbering creates a minor navigational disruption, the layout largely compensates by maintaining a recognizable visual sequence. This case supports Kress and van Leeuwen's (2021) distinction between narrative and conceptual visual processes and provides empirical evidence that action vectors are particularly important in procedural genres. It also reinforces O'Halloran's (2008) theoretical claim that combining linguistic and visual forms may expand ideational meaning rather than simply duplicate information.

Another significant contribution concerns the students' use of cultural resources. Dayak motifs, hornbills, orangutans, traditional houses, warriors, dancers, and ornamental patterns occur repeatedly across the corpus. These elements show that digital multimodal composing is not culturally neutral. The students appropriate the global technological affordances of Canva while embedding local semiotic resources drawn from their Central Kalimantan environment. This finding accords with social semiotic perspectives that treat meaning-making as socially and culturally situated rather than as the mechanical application of universal sign systems (Bezemer & Kress, 2016; Van Leeuwen, 2006). It also extends research such as Sugianto et al. (2022), which examined multimodal meaning in published EFL textbook materials, because the present study focuses on learners as active designers of multimodal texts rather than recipients of professionally designed materials. The "Localization Strategy" identified here therefore represents a form of learner agency: students recontextualize a standardized digital genre through culturally recognizable imagery. Nevertheless, the findings also show that cultural representation

must be balanced with communicative functionality. In several artifacts, extensive decorative motifs contributed to visual density and competed with the instructional sequence for reader attention. Cultural localization is therefore most effective when it strengthens identity without weakening informational hierarchy.

The finding of Hybrid Modality further demonstrates students' strategic use of different levels of visual realism. Several infographics combine realistic photographs of finished products with simplified icons for intermediate processes. This combination can be interpreted as an attempt to exploit different modal affordances: realistic photography establishes the desired end-state with high representational specificity, while simplified graphics conserve space and support rapid identification. Such practices resonate with multimodal theories emphasizing that individual modes possess different representational potentials and constraints (Kress & van Leeuwen, 2021; Cope & Kalantzis, 2020). However, hybrid modality becomes problematic when the simplified icons lack sufficient action information. The finding therefore indicates that selecting multimodal resources should be guided not only by aesthetic consistency but by the experiential meaning each resource is expected to realize.

Importantly, the study also demonstrates that students' multimodal choices cannot be interpreted independently of their socio-technical environment. Interviews and classroom observations indicate that unstable Wi-Fi connectivity within the remote plantation context sometimes encouraged students to select readily available templates or basic icons rather than search for more specific visual representations. This finding adds a contextual dimension often underrepresented in multimodal discourse research. Prior and Shipka (2018) conceptualize composing as distributed across people, tools, spaces, and material conditions; the present results demonstrate concretely how infrastructure becomes part of this distributed meaning-making system. Students' semiotic choices therefore cannot always be interpreted simply as individual competence or preference. Some apparent design weaknesses may partly reflect constrained access to technological resources.

CONCLUSION

This study investigated how ninth-grade EFL students realized experiential meaning through linguistic, visual, spatial, and cultural resources in digital procedure infographics created within a remote palm oil plantation school in Central Kalimantan. The findings demonstrate that experiential meaning was most systematically realized in the verbal mode through the dominance of Material Processes, typically expressed through imperative action verbs acting upon Goals, with Circumstances adding information about place, manner, extent, and condition. Several artifacts additionally exhibited a Relational Shift at their closure, linguistically marking the movement from processes of transformation to states of completion. Visual realization was considerably more variable. While most students employed static Conceptual Structures to identify materials, products, or cultural symbols, a smaller number successfully deployed Narrative Processes and action vectors to represent the physical actions encoded in their verbal instructions. The relationship between these modes produced both inter-semiotic complementarity and breakdown. The most consequential breakdown, conceptualized in this study as the Static Paradox, occurred when dynamic verbal instructions were accompanied by static images that failed to visualize movement. Conversely, the hand-washing infographic demonstrated how explicit action vectors, structured panels, and coherent reading paths could establish strong verbal–visual congruence.

Beyond verbal–visual alignment, this study establishes that multimodal meaning-making is deeply shaped by cultural identity and material conditions. Students localized globally

recognizable infographic conventions through Dayak patterns, regional wildlife, traditional houses, dancers, warriors, and other Borneo-related visual resources, demonstrating that learner-created digital texts can become spaces for the negotiation and representation of local identity. At the same time, excessive cultural ornamentation sometimes increased visual density, while unstable internet connectivity constrained students' access to appropriate graphics and contributed to dependence on static, pre-packaged visual resources. Consequently, multimodal competence should not be understood solely as an individual cognitive or linguistic ability but as a socially situated capacity shaped by cultural knowledge, pedagogical scaffolding, digital resources, and technological infrastructure. The study consequently contributes to ELT and multimodal discourse research by connecting SFL transitivity, Visual Grammar, intermodal relations, local cultural representation, and rural technological realities within one analytical framework. Pedagogically, teachers should move beyond assessing only linguistic accuracy and visual attractiveness toward explicitly teaching how actions, participants, circumstances, images, vectors, layout, and cultural resources work together to construct experiential meaning.

FUNDING

This research was entirely self-funded by the authors. The study received no specific grant, financial assistance, or sponsorship from any funding agency, institution, or third party in the public, commercial, or not-for-profit sectors.

INFORMED CONSENT STATEMENT

Participation in this study was entirely voluntary. Because the student participants in this study were minors aged fourteen to fifteen years old, a multi-tiered ethical clearance protocol was strictly implemented. Prior to the research, formal institutional authorization was obtained from the Principal of SMP Eka Tjipta Semilar and the managing Eka Tjipta Foundation (Sinar Mas Group). Furthermore, active written informed consent was secured from the parents or legal guardians of all participating students, and written assent was obtained directly from the ninth-grade students.

DATA AVAILABILITY STATEMENT

The raw qualitative dataset including the unedited student infographics and unredacted interview transcripts cannot be made publicly available due to strict ethical obligations safeguarding minors and confidentiality agreements with the Eka Tjipta Foundation regarding the plantation operational area. However, anonymized data records including the standardized codebooks, inter-semiotic coding matrices, and de-identified selections of interview transcripts are available under controlled access. Interested parties may request access to these resources by contacting the corresponding author.

DECLARATION IN THE USE OF AI

In order to improve the clarity and readability of the manuscript, artificial intelligence (AI) tools were used appropriately. Additionally, the content in the manuscript has been thoroughly reviewed, revised, and refined by the researchers.

ACKNOWLEDGEMENT

The researchers would like to express their sincere gratitude to colleagues for their invaluable guidance, expert supervision, and continuous encouragement throughout the development of this research. We also extend our appreciation to the Master Program of English Education at Universitas Ahmad Dahlan for providing the institutional support necessary for this study. Special thanks are due to the principal and staff of SMP Eka Tjipta Semilar, as well as the Eka

Tjipta Foundation (Sinar Mas Group), for granting formal authorization and facilitating the data collection process within the plantation operational area. Finally, we are deeply grateful to the ninth-grade students of the 2025/2026 academic year for their participation and for contributing the creative digital infographics that served as the primary data for this multimodal analysis.

REFERENCES

- Bezemer, J., & Kress, G. (2016). *Multimodality, learning and communication: a semiotic frame*. Routledge. <https://doi.org/10.4324/9781315687537>
- Bloor, T., & Bloor, M. (2020). *The functional analysis of English: A Hallidayan approach* (4th ed.). Routledge.
- Bowen, C. J. (2017). *Grammar of the edit* (4th ed.). Routledge. <https://doi.org/10.4324/9781315208404>
- Brett, D., Fahey, R., Feez, S., Spinks, S., & Yallop, C. (2023). *Functional grammar: An introduction for the real world* (5th ed.). Routledge. <https://doi.org/10.4324/9781003283911>
- Butt, D., Moore, R., Berguson, A., Matthiessen, C. M. I. M., & Spinks, M. (2023). *Using functional grammar: An explorer's guide* (4th ed.). Routledge.
- Cope, B., & Kalantzis, M. (2020). *Making sense: Reference, agency and structure in a grammar of multimodal meaning*. Cambridge University Press. <https://doi.org/10.1017/9781316459645>
- Corderio, C. M. (2018). Using systemic functional linguistics as method in identifying semogenic strategies in intercultural communication. *Journal of Intercultural Communication*, 18(1), 1–15.
- Creswell, J. W., & Creswell, J. D. (2018). *Research design: Qualitative, quantitative, and mixed methods approach* (5th ed.). SAGE Publications.
- Creswell, J. W., & Poth, C. N. (2018). *Qualitative inquiry and research design: Choosing among five approaches* (4th ed.). SAGE publications.
- Feist, J. (2022). *Significance in language: A theory of semantics*. Routledge. <https://doi.org/10.4324/9781003259381>
- Halliday, M. A. K. (2014). *Language as social semiotic: The social interpretation of language and meaning*. University Park Press. (Original work published 1978)
- Halliday, M. A. K., & Matthiessen, C. M. I. M. (2014). *Halliday's introduction to functional grammar* (4th edition). Routledge. <https://doi.org/10.4324/9780203783771>
- Huang, X., O'Halloran, K.L., & Wignell, P. (2021). *Multimodal analysis of classroom discourse and interactions*. Routledge.
- Indrayani, R. (2018). *A multimodal discourse analysis of the images of G-Dragon reflected on his Instagram account* (Unpublished bachelor's thesis). Universitas Brawijaya.
- Jewitt, C. (2017). *Introducing multimodality*. Routledge. <https://doi.org/10.4324/9781315638027>
- Knapp, P., & Watkins, M. (2005). *Genre, text, grammar: Technologies for teaching and assessing writing*. UNSW Press.
- Kolb, A. Y., & Kolb, D. A. (2018). Eight important things to know about the experiential learning cycle. *Australian Educational Leader*, 40(3), 8 – 14.
- Kress, G., & van Leeuwen, T. (2021). *Reading images: The grammar of visual design* (3rd ed.). Routledge. <https://doi.org/10.4324/9781003099857>
- Kyaw, A. M. (2024). Comparison between the modality analyses of two Ellen Talk Shows. *Parole: Journal of Linguistics and Education*, 14(1), 66 – 73. <https://doi.org/10.14710/parole.v14i1.66-73>

- Liu, Y., & O'Halloran, K. L. (2019). Intersemiotic management of meaning in visual communication. *Sign Systems Studies*, 47 (3/4), 478 – 507. <https://doi.org/10.12697/SSS.2019.47.3-4.07>
- Marian, V. (2024). *The power of language: How the codes we use to think, speak, and live transform our minds*. Penguin Random House.
- Martin, J. R., & Knox, T. (2022). *Systemic functional linguistic: A complete guide to the theory and the tools*. Routledge. <https://doi.org/10.4324/9781003181460>
- Martin, J. R. (2024). *Context as a system of resources realized through linguistic choices*. Cambridge University Press. <https://www.scribd.com/document/486278895/38-Working-with-Discourse-Martin-and-Rose>
- Megah S., S. I. (2019). A systemic functional linguistic analysis of the transitivity in Jokowi's new vision speech. *Cahaya Pendidikan*, 4(2), 51 – 61. <https://doi.org/10.37074/cahaya.v4i2.2225>
- Miles, M. B., Huberman, A. M., & Saldana, J. (2014). *Qualitative data analysis: A methods sourcebook* (3rd ed.). SAGE Publications.
- Mulki R., N. M., Sahril, & Sunra, L. (2022). Analyzing the students' errors in writing procedure text. *Journal of Excellence in English Language Education*, 1(1), 12-25.
- Noble, H., & Heale, R. (2019). Triangulation in research, with example. *Evidence-Based Nursing*, 22(3), p. 67-68. <https://doi.org/10.1136/ebnurs-2019-103145>
- O'Halloran, K. L. (2008). Systemic functional-multimodal discourse analysis (SF-MDA): Constructing ideational meaning using language and visual imagery. *Visual Communication*, 7(4), 443-501. <https://doi.org/10.1177/1470357208096210>
- Pauwels, L., & Mannay, D. (Eds.). (2019). *The SAGE handbook of visual research methods* (2nd ed.). SAGE Publications. <https://doi.org/10.4135/9781529721485>
- Prior, P., & Shipka, J. (2018). *Writing – in – the – world: Distributing creative agency* (2nd ed.). Routledge. <https://doi.org/10.4324/9781315545677>
- Putri, S. J. (2024). *A survey of multimodal composition in creative writing class in english education study program of Sriwijaya University* (Unpublished bachelor's thesis). Sriwijaya University.
- Sadhu, D., & Ghasani, T. D. (2023). Kemandirian belajar siswa dalam pembelajaran menggunakan kurikulum merdeka. *Plusminus: Jurnal Pendidikan Matematika*, 3(2), 241-254. <https://doi.org/10.31980/plusminus.v3i2.1346>
- Sugianto, A., Prasetyo, I. A., & Asti, W. (2022). 'Is the picture worth a thousand words?': The interpersonal meanings of a dialogue in an EFL textbooks. *JOLLT Journal of Languages and Language Teaching*, 10(2), 187 – 198. <https://doi.org/10.33394/jollt.v%vi%i.4510>
- Sweller, J. (2024). Cognitive load theory and the curriculum. In M. Young & D. Lambert (Eds.). *Research handbook on curriculum and education* (pp. 155–166). Edward Elgar Publishing. <https://doi.org/10.4337/9781802208542.00017>
- Tarigan, H. G. (2020). *Pengajaran semantik*. Angkasa
- Van Leeuwen, T. (2006). *Introducing social semiotics*. Routledge
- Vicente, A. (2018). Polysemy and word meaning: An account of lexical meaning for different kinds of content word. *Philosophical Studies*, 175(2), 347–368. <https://doi.org/10.1007/s11098-017-0900-y>
- Zhang, D. (2018). *Multimodal discourse analysis: A systemic functional synthetic framework* (2nd ed.). Higher Education Press.