

---

## IMPROVING STUDENTS' SPEAKING SKILLS THROUGH CARTOON MEDIA: A CLASSROOM ACTION RESEARCH AT AN INDONESIAN ELEMENTARY SCHOOL

Oleh

Jaisy Yasmin Alphariyadi<sup>1</sup>, Lusiana Suciati Dewi<sup>2</sup>, Muhammad Anjar Nugraha<sup>3</sup>

<sup>1,2,3</sup>English Language Education Study Program, Faculty of Teacher Training and Education, Universitas Subang

Email: <sup>1</sup>[jaisyyasmin12@gmail.com](mailto:jaisyyasmin12@gmail.com), <sup>2</sup>[lusianasd@unsub.ac.id](mailto:lusianasd@unsub.ac.id),

<sup>3</sup>[Anjarnugraha@unsub.ac.id](mailto:Anjarnugraha@unsub.ac.id)

---

### Article History:

Received: 01-06-2026

Revised: 26-06-2026

Accepted: 04-07-2026

### Keywords:

Cartoon Media;  
Speaking Skills;  
Classroom Action  
Research; EFL;  
Elementary School

**Abstract:** *This study investigates the effectiveness of cartoon media in improving the English-speaking skills of sixth-grade students at an elementary school in Subang, Indonesia. Employing a Classroom Action Research (PTK) design following the Kemmis and McTaggart model, the research was conducted across two cycles, each comprising planning, action, observation, and reflection phases. Data were collected through pre- and post-tests assessed via a speaking rubric (pronunciation, fluency, vocabulary, and grammar), classroom field notes, semi-structured interviews, and documentation. In Cycle 1, despite increased classroom participation, the post-test mean score declined from 60 to 54, attributable to unfamiliar vocabulary and persistent speaking anxiety. Informed by this reflection, Cycle 2 incorporated vocabulary pre-teaching and dubbing activities, resulting in a substantial improvement: the post-test mean rose to 64, with 76% of students reaching the mastery threshold of 70 or above, surpassing the predetermined 75% success criterion. Qualitative data corroborated these gains, with students reporting reduced anxiety and increased enjoyment when learning through cartoon-based activities. The findings support Krashen's affective filter hypothesis and Mayer's Cognitive Theory of Multimedia Learning, suggesting that cartoon media can serve as a low-cost, engaging, and pedagogically effective tool for developing oral communicative competence among young EFL learners.*

---

## INTRODUCTION

Speaking is widely regarded as one of the most challenging language skills to master, particularly within English as a Foreign Language (EFL) contexts where learners have limited exposure to authentic spoken input. Speaking requires the simultaneous coordination of vocabulary retrieval, grammatical encoding, pronunciation, and real-time processing, a demand that places considerable cognitive load on learners, especially those at the elementary level who are still consolidating foundational literacy and oral skills in their first language. Many Indonesian elementary school students struggle to speak English

comfortably, often due to limited vocabulary, lack of confidence, and elevated anxiety when performing in front of peers and teachers. This reluctance is compounded by limited exposure to authentic spoken English outside the classroom and by the comparatively high cost of structured speaking courses, which remain inaccessible to many learners and families.

Speaking ability is widely recognized as foundational to learners' broader communicative competence and future academic and professional success. Individuals with stronger oral communication skills tend to build more effective interpersonal relationships and access a wider range of professional opportunities. Within the framework of twenty-first century skills, learners are increasingly expected to develop competencies across four broad domains: ways of thinking, including creative and critical thinking; ways of working, including communication and collaboration; tools for work, including information and digital literacy; and skills for living, including social responsibility and civic participation (Binkley et al., 2012). This framework positions oral communication not merely as a discrete linguistic skill but as a foundation for participation in an increasingly globalized and digitally mediated society, lending particular urgency to the search for effective, accessible speaking pedagogies.

Despite this recognized importance, speaking instruction in many Indonesian EFL classrooms continues to rely heavily on rote-based, teacher-centered methods that provide insufficient opportunities for meaningful oral practice. Overcrowded classrooms and limited access to authentic communicative materials further constrain opportunities for students to engage in spontaneous, low-anxiety speaking practice. These structural and pedagogical limitations frequently result in students who, despite years of formal English instruction, remain unable to describe, explain, or discuss everyday topics with confidence or fluency.

In response to these challenges, multimedia-based instruction has emerged as a promising pedagogical avenue. Prior studies have demonstrated that visual media, including cartoons and picture-cued materials, can reduce speech errors, increase motivation, and improve speaking outcomes among young learners. Cartoons and picture-cued media have been shown to improve primary school students' speaking abilities while reducing speech errors and increasing motivation (Duran et al., 2021). Audio-visual media generated through applications such as Cartoon Story Maker have similarly proven effective in improving fifth-grade students' speaking performance (Sandi et al., 2023), while cartoon imagery has been successfully applied to develop speaking skills among learners of Indonesian as a foreign language at the beginner level, with particular gains in vocabulary accuracy, sentence construction, and pronunciation (Karina et al., 2023). Collectively, these studies suggest that visual, narrative-based media can render language learning more engaging, increase active participation, and elicit positive affective responses from learners, positioning cartoon media as a promising alternative to conventional, teacher-centered speaking instruction.

Despite this growing evidence base, comparatively little research has examined the use of cartoon media within an iterative, cyclical Classroom Action Research framework in Indonesian elementary classrooms specifically, where teachers must continuously adapt instruction in response to ongoing classroom evidence rather than relying on a single, fixed intervention design. Action research offers a particularly suitable methodology for this purpose, as its cyclical structure of planning, action, observation, and reflection allows instructional strategies to be iteratively refined in direct response to observed student

outcomes, rather than assuming uniform effectiveness across a static intervention period. The present study therefore investigates the extent to which cartoon media, implemented through a two-cycle Classroom Action Research design, can improve sixth-grade students' speaking skills, while also examining students' affective and motivational responses to the intervention across successive cycles of refinement.

Specifically, this study addresses two research questions: (1) Can cartoon media improve students' speaking skills among sixth-grade learners at an Indonesian elementary school? and (2) What are students' responses to learning speaking through cartoon media, particularly with respect to motivation, engagement, and speaking anxiety? By situating cartoon media within an explicitly iterative research design, this study aims to contribute not only evidence of cartoon media's instructional effectiveness, but also practical insight into the specific pedagogical refinements, such as vocabulary scaffolding and dubbing-based practice, that are necessary to realize its full potential in real classroom conditions.

## LITERATURE REVIEW

### *Speaking Skills in EFL Classrooms*

Speaking is a complex, multifaceted skill requiring the simultaneous coordination of lexical retrieval, grammatical encoding, and real-time pronunciation, which places considerable cognitive load on language learners, particularly young learners still developing foundational linguistic competence (Bygate, 1987). Unlike receptive skills such as listening or reading, speaking demands immediate, real-time production under social observation, a condition that introduces both cognitive and affective pressures absent from most other language skills. Teaching speaking to young EFL learners therefore presents challenges across cognitive, affective, and performance domains, as systematically categorized by Thornbury (2005). Cognitively, learners must draw on limited lexical resources, often imprecise grammatical knowledge, and developing phonological awareness, frequently while also managing first-language interference and characteristically brief attention spans typical of elementary-aged learners.

Affective barriers, particularly anxiety and shyness, generate an elevated "affective filter" that constrains spontaneous oral production (Krashen, 1982). This barrier is especially pronounced in classroom settings, where fear of peer judgment frequently inhibits even moderately proficient students from attempting spoken output (Young, 1991). Such anxiety is rarely distributed evenly across a classroom; rather, it tends to be most acute precisely among the students who would benefit most from additional speaking practice, creating a self-reinforcing cycle in which anxious learners avoid practice opportunities that might otherwise reduce their anxiety over time. Performance-related challenges, including limited fluency, hesitations, and disfluencies, often emerge as the visible surface manifestation of these underlying cognitive and affective difficulties, rather than as discrete problems in their own right.

These cognitive and affective difficulties are often compounded by first-language interference, as phonemic distinctions absent in students' native languages frequently produce persistent mispronunciation and hesitation in English discourse (Deterding & Kirkpatrick, 2006). In the Indonesian context specifically, students must additionally navigate phonological and structural differences between English, Bahasa Indonesia, and

regional mother tongues, each of which may introduce distinct interference patterns. Traditional, teacher-centered pedagogical approaches that remain prevalent in Indonesian primary classrooms offer limited opportunities for authentic communicative interaction, further constraining the development of functional oral competence (Lisnawati et al., 2017). Overcrowded classroom conditions and a paucity of authentic practice opportunities frequently precipitate motivational decline among learners, who may come to view speaking practice as a source of stress rather than growth.

These pervasive constraints collectively underscore the need for innovative, low-anxiety instructional interventions capable of delivering contextualized linguistic input while actively mitigating affective barriers to speaking. Animated media, including cartoons, have demonstrated considerable efficacy in sustaining learner engagement and rendering target language input comprehensible through visual contextualization, narrative repetition, and affectively accessible content, positioning cartoon-based instruction as a strategically appropriate response to the specific challenges identified above.

### ***Multimedia-Based Instruction in Speaking Development***

Multimedia instruction, defined as the integrated use of two or more media types, such as text, audio, graphics, animation, and video, within a single instructional activity (Ariyanto et al., 2018; Zhen, 2016), has been shown to enhance learner motivation, engagement, and communicative competence (Gilakjani, 2012). Theoretical support for multimedia-based language instruction is provided by several complementary frameworks. Mayer's (2014) Cognitive Theory of Multimedia Learning posits that the combined use of visual and auditory channels facilitates more effective cognitive processing than single-channel input alone, on the basis that working memory possesses separate, complementary channels for processing visual and auditory information, the simultaneous engagement of which reduces the processing burden carried by either channel individually. Paivio's (1986) Dual Coding Theory similarly suggests that information encoded through both visual and verbal channels is retained and recalled more effectively than information encoded through a single channel, lending further theoretical support to the integration of visual cartoon content with spoken dialogue.

Within speaking instruction specifically, multimedia can be pedagogically sequenced through pre-speaking, speaking, and post-speaking stages (Abidin, 2015). The pre-speaking stage encompasses preparatory activities such as identifying the topic, establishing communicative purpose, and activating relevant vocabulary and background knowledge. The speaking stage itself involves guided practice and structured oral production, while the post-speaking stage focuses on reflection, feedback, and consolidation of the speaking experience. This three-stage sequencing provides a structured scaffold through which multimedia content, including cartoon clips, can be deliberately integrated to support progressive oral skill development, rather than being deployed merely as an engaging but pedagogically unstructured supplement to instruction.

Empirical research further substantiates the pedagogical value of multimedia in speaking instruction. Multimedia teaching has been shown to build meaningful interaction between teachers and students, particularly in language learning contexts where its principal aim is to develop listening and speaking competence (Huang et al., 2008). Within such instructional designs, the teacher's role shifts from sole transmitter of knowledge toward

facilitator, while multimedia content itself contributes to a more dynamic and engaging classroom environment. Multimedia instruction is similarly noted for its flexibility, as it is not confined to a fixed classroom setting or time, and for the collaborative opportunities it affords students working with shared audio-visual content. Effective multimedia use, however, requires adherence to certain guiding principles: content must be carefully managed to avoid cognitive overload, learners' prior knowledge should be activated before new multimedia content is introduced, and opportunities for active learner interaction and timely feedback should be deliberately built into the instructional design (Gilakjani, 2012). These principles directly inform the present study's instructional design, particularly the decision to introduce vocabulary pre-teaching prior to cartoon viewing in Cycle 2.

### ***Cartoon Media as a Speaking Instruction Tool***

Cartoon media, encompassing animated television programs, films, and short instructional clips, represents a particularly accessible and cost-effective multimedia format for EFL speaking instruction. Cartoons expose learners to natural, conversational dialogue that models tone, intonation, stress, and rhythm in ways that learners can directly imitate and internalize, thereby supporting the development of pronunciation and fluency within a low-pressure, engaging environment. Beyond their linguistic function, cartoons situate language within meaningful real-life contexts, illustrating how specific words and phrases are used to express emotions, ask questions, and navigate everyday social situations, thereby strengthening learners' practical communicative competence alongside their formal linguistic knowledge.

Among the various pedagogical activities associated with cartoon-based instruction, voiceover dubbing, in which students view a muted cartoon clip and construct their own spoken dialogue to match the visual narrative, has been identified as particularly effective in developing pronunciation, intonation, and expressive oral delivery. Dubbing tasks require learners to think quickly, select contextually appropriate language, and align their spoken output with the emotional and narrative tone of the animation, thereby integrating cognitive, linguistic, and performative dimensions of speaking practice within a single engaging activity. Critically, because dubbing activities are framed as a playful, creative exercise rather than a formal speaking assessment, they appear to substantially reduce the anxiety typically associated with spontaneous oral performance, rendering them especially well-suited to anxious or reluctant speakers.

Notwithstanding these advantages, cartoon-based instruction is not without limitations. The language featured in cartoons, particularly those designed for child audiences, is frequently simplified and may therefore provide insufficient exposure to the more complex syntactic structures required for formal or academic communication. Cartoons may also embed culturally specific humor or references that prove confusing to learners unfamiliar with the relevant cultural context, while an over-reliance on visual cues risks limiting learners' independent development of listening comprehension absent visual support. These limitations indicate that cartoon media is most appropriately deployed not as a standalone intervention, but as one component within a broader, carefully scaffolded instructional design, a principle that directly informed the iterative refinements introduced across the two cycles of the present study.

### ***Related Studies and Research Gap***

Several recent studies have investigated the application of cartoon and related visual media to speaking instruction across varied educational contexts, collectively informing the design of the present study. Duran et al. (2021) examined the use of cartoons within a multigrade primary classroom setting and found that cartoon-based instruction was associated with measurable reductions in speech errors alongside increased learner motivation, suggesting that even within the structurally complex conditions of multigrade instruction, cartoon media can support meaningful gains in oral production. Sandi et al. (2023) investigated the use of audio-visual media generated through the Cartoon Story Maker application among fifth-grade elementary students, reporting clear improvements in speaking skills following the intervention and underscoring the pedagogical value of teacher-generated, customizable cartoon content as opposed to commercially produced animation alone.

Karina et al. (2023) extended this line of inquiry to the context of Indonesian as a foreign language instruction at the beginner proficiency level, demonstrating that cartoon imagery could be effectively applied to develop vocabulary accuracy, sentence construction, and pronunciation among adult foreign-language learners, suggesting that the pedagogical benefits associated with cartoon media are not necessarily confined to child learners or to English-language instruction specifically. Taken collectively, these studies converge on a consistent finding: cartoon and cartoon-adjacent visual media reliably increase learner engagement and participation, and are frequently, though not invariably, associated with corresponding improvements in measurable speaking outcomes.

Despite this convergence, the existing literature exhibits a notable methodological gap that the present study is positioned to address. The studies reviewed above predominantly employ single-intervention or quasi-experimental designs, comparing outcomes before and after a single, relatively fixed period of cartoon-based instruction. Comparatively few studies have examined cartoon media through an explicitly iterative, cyclical action research framework, in which instructional design is deliberately revised between cycles in direct response to observed classroom outcomes. This represents an important gap, as a single-intervention design cannot capture the possibility, demonstrated clearly in the present study, that cartoon media's effectiveness may depend critically on surrounding pedagogical scaffolding that is absent in an initial implementation but can be identified and incorporated through subsequent cycles of reflective practice. The present study addresses this gap directly, contributing not only further evidence of cartoon media's general effectiveness, but a specific, replicable account of how that effectiveness can be incrementally realized through structured reflection and refinement.

## **METHODOLOGY**

### ***Research Design***

This study employed a Classroom Action Research (PTK) design following the Kemmis and McTaggart (1988) model, a methodology widely adopted in educational research for its capacity to generate practical, classroom-grounded knowledge through cycles of deliberate intervention and reflective evaluation. The design comprised two cycles, each consisting of four interrelated phases: planning, in which lesson plans, instructional materials, and

assessment instruments were prepared; action, in which the planned intervention was implemented in the classroom; observation, in which data were systematically collected on student performance and behavior; and reflection, in which the collected data were analyzed to evaluate the intervention's effectiveness and inform refinements for the subsequent cycle. This cyclical structure is particularly well-suited to classroom contexts, as it allows the researcher to respond directly to emerging evidence rather than committing to a single, fixed instructional design for the duration of the study.

### ***Research Setting and Participants***

The research was conducted at SDN Ir. H. Juanda, a public elementary school located in Subang, West Java, Indonesia. The population of the study comprised all sixth-grade students at the school. Given the relatively small and manageable size of the population, no formal sampling procedure was applied; rather, all sixth-grade students were included as research participants, consistent with a total sampling approach. Thirty-one students participated in Cycle 1, while twenty-nine students participated in Cycle 2, the difference attributable to attendance variation across the research period rather than to any change in the underlying study population. The sixth grade was selected as the research site on the basis of preliminary classroom observation, which identified speaking anxiety and limited oral fluency as particularly prominent challenges among students at this level. The study was conducted over an eight-week period between August and December, encompassing an initial planning phase, a pre-test, two complete instructional cycles, and corresponding reflection phases.

### ***Data Collection Instruments***

Data were collected through four triangulated instruments, integrating both quantitative and qualitative sources to provide a comprehensive evidentiary basis for evaluating the intervention. First, a speaking performance rubric assessed pronunciation (30%), fluency (30%), vocabulary (20%), and grammar (20%) on a standardized 0–100 scale, applied during pre- and post-tests in both cycles. The relative weighting of pronunciation and fluency reflects the study's pedagogical emphasis on these two domains as the most directly observable outcomes of cartoon-based instruction. A success criterion was established prior to the commencement of the study, requiring 75% of students to achieve a minimum score of 70, corresponding to a "Good" rating, for the intervention to be deemed successful and the action research concluded.

Second, classroom field notes captured real-time observations of student participation, pronunciation accuracy during speaking activities, and affective indicators such as hesitation, reluctance to speak, and visible signs of self-consciousness. Rather than employing a fixed observational checklist, the researcher adopted a naturalistic, descriptive field-note approach consistent with the reflective orientation of the Kemmis and McTaggart model, recording observations as they emerged across each instructional session. The researcher served as the primary observer throughout both cycles; to mitigate potential observer bias and strengthen the credibility of the observational data, these field notes were supplemented by the concurrent observations of a non-participant collaborator.

Third, semi-structured interviews were conducted with three purposively selected students per cycle, representing low, moderate, and high performance levels respectively, in order to capture a breadth of experiential perspectives reflective of the full range of learner proficiency within the classroom. The semi-structured format was adopted for its capacity to

balance consistency across interviews with sufficient flexibility to allow the researcher to pursue emergent themes and encourage participants to elaborate on their experiences in their own terms. Interview questions were designed to elicit affective, motivational, and metacognitive dimensions of the learning experience, including students' perceived changes in confidence, anxiety, and vocabulary development across the intervention.

Fourth, documentation was collected as a supplementary data source, comprising lesson plans, pre- and post-test score sheets, student dubbing recordings, and classroom photographs obtained in accordance with informed consent protocols. The comparative analysis of student dubbing recordings across the two cycles, in particular, enabled detailed assessment of vocabulary usage, pronunciation accuracy, and overall fluency, revealing patterns of linguistic development that corroborated the quantitative test results and enriched the interpretive depth of the study's findings.

### ***Procedure***

In Cycle 1, instructional planning centered on short clips, approximately sixteen to seventeen minutes in duration, drawn from the animated film *Cars*, selected for their appropriateness to sixth-grade proficiency levels. Across four instructional meetings, students completed a pre-test requiring them to describe simple images, followed by cartoon screening, vocabulary and plot discussion, brief role-play of key dialogues, guided pronunciation practice through imitation of the researcher's modeled examples, pair discussion of the narrative's moral themes, and a post-test centered on a dubbing task. In the final meeting of the cycle, students submitted completed dubbing projects, which were collectively reviewed for pronunciation and other errors.

Reflection on Cycle 1 outcomes revealed that, despite measurable increases in classroom participation following cartoon exposure, students' post-test performance had not improved relative to the pre-test, with key limitations identified as inadequate pre-viewing of target vocabulary and persistent student absenteeism. These findings directly informed the refinements introduced in Cycle 2, which incorporated structured vocabulary pre-teaching, targeting ten key words per cartoon clip, alongside dubbing-focused practice activities and attendance reminders disseminated to parents via WhatsApp. Cycle 2 instruction centered on clips from the animated film *Ice Age*, selected for their adventure theme, and was delivered across three instructional meetings following a structure analogous to that of Cycle 1, culminating in a final dubbing-based post-test and group-based review of pronunciation errors.

### ***Data Analysis Technique***

Data analysis followed a mixed-methods approach integrating quantitative and qualitative procedures consistent with the study's triangulated data collection framework. Quantitative speaking scores obtained from the pre-test and post-test in each cycle were analyzed descriptively, with the mean score for each test calculated as the total score of all students divided by the number of students. The mastery percentage was subsequently calculated as the proportion of students achieving a score of 70 or above relative to the total number of students, expressed as a percentage, in order to assess whether the predetermined 75% success criterion had been met at the conclusion of each cycle. Qualitative data obtained from field notes and interviews were analyzed thematically, with recurring patterns relating to student participation, hesitation, anxiety, motivation, and

engagement with cartoon-based activities identified and grouped into thematic categories. These qualitative findings were subsequently cross-referenced with the quantitative results to produce a triangulated interpretation of the intervention's overall effectiveness across both cycles.

It should be acknowledged that, as with most classroom-based action research, this study is subject to certain methodological constraints inherent to its design. The absence of a control group means that observed improvements between Cycle 1 and Cycle 2 cannot be attributed to the instructional refinements alone with the same certainty afforded by an experimental design, since some portion of the observed gain may be attributable to general familiarity with the testing format or to natural variation in student readiness across the research period. Similarly, the relatively brief duration of the intervention, spanning a single academic term, limits the extent to which conclusions can be drawn regarding the durability of the speaking gains documented at the conclusion of Cycle 2. These constraints are typical of action research conducted within real classroom conditions, where the priority of generating timely, practically actionable insight for the participating classroom must be balanced against the more stringent methodological controls characteristic of experimental designs, and they are addressed further in the discussion of limitations below.

## FINDINGS AND DISCUSSION

### *Cycle 1: Implementation and Outcomes*

During Cycle 1, the pre-test produced a mean score of 60, calculated from a total score of 1,859 across 31 students, with a minimum score of 0 and a maximum score of 90. The post-test, by contrast, produced a mean score of 54, calculated from a total score of 1,677, with a minimum of 0 and a maximum of 90. This outcome represented a decline rather than the anticipated improvement, a finding that required careful interpretation in light of the qualitative data collected alongside the quantitative scores. Classification of student performance revealed that the proportion of students scoring in the "Good" range increased substantially from pre-test to post-test, alongside a marked increase in the proportion of students whose results fell in the "Excellent" range, even as the overall mean declined, suggesting that gains were concentrated among a subset of students while others, particularly those affected by absenteeism or persistent anxiety, did not progress at the same rate.

Field notes revealed that approximately 70% of students exhibited passive behavior during the pre-test, characterized by nervousness and prolonged pauses attributable to shyness, consistent with the affective barriers to speaking documented extensively in the literature reviewed above. Following cartoon screening, however, classroom participation increased by approximately 40%, with students demonstrating enthusiastic mimicry of character gestures and dialogue, and visual aids appearing to facilitate retention of key vocabulary items, contributing to an estimated 25% reduction in pronunciation errors during guided practice. Observable indicators of growing engagement included collective laughter during cartoon screening and a marked increase in spontaneous hand-raising, rising from approximately five to fifteen instances across the observed sessions, illustrating the capacity of cartoon media to enhance student confidence through engaging, contextually grounded immersion, notwithstanding the disruptive effect of ten student absences during

the testing period.

Reflection on these outcomes indicated that while cartoon-based activities successfully enhanced engagement and motivation, as corroborated by student interview responses describing the activities as "fun" and "easier to speak," the post-test score decline was attributable to a combination of unfamiliar vocabulary encountered during the dubbing task, persistent initial shyness, and the rigor of the scoring rubric applied to the post-test, compounded by the disruptive effect of student absenteeism. These findings indicated that the instructional design required targeted refinement, specifically with respect to vocabulary preparation, before the desired learning outcomes could be more fully achieved. Accordingly, the research proceeded to Cycle 2, incorporating targeted vocabulary pre-teaching, structured dubbing exercises, and attendance reminders disseminated to parents.

### ***Cycle 2: Implementation and Outcomes***

Cycle 2 produced a pre-test mean score of 50, calculated from a total score of 1,561 across 31 students, with a minimum score of 0 and a maximum score of 85, and a post-test mean score of 64, calculated from a substantially higher total score, with a minimum of 0 and a maximum of 95. This represented a 14-point improvement from pre-test to post-test, a marked reversal of the pattern observed in Cycle 1. Classification of post-test performance indicated that 76% of students achieved mastery, comprising 52% of students at the "Good" level and 24% at the "Excellent" level, thereby surpassing the study's predetermined 75% mastery criterion and meeting the threshold required to conclude the action research.

Field notes recorded a marked increase in active classroom participation, rising to approximately 90% compared to 60% in the corresponding phase of Cycle 1, alongside an estimated 30% reduction in speech pauses during dubbing exercises, attributable to the enhanced fluency afforded by the cycle's vocabulary pre-teaching and structured practice activities. Dubbing exercises in particular generated notable student enthusiasm through mimicry of the Ice Age characters' intonation patterns, while also promoting more natural gestural expression during speech production. Despite seven student absences during the cycle, overall classroom engagement remained robust, with field notes recording universal smiling during speech production, an observation interpreted as indicative of meaningfully reduced speaking anxiety relative to Cycle 1. Interview responses gathered at the conclusion of Cycle 2 further corroborated these quantitative gains, with students describing the activities as making speaking "easier" thanks to the "fun visual" elements of the cartoon-based instruction.

Table 1 summarizes the comparative pre-test and post-test outcomes across both cycles, while Table 2 presents the corresponding classification of post-test performance levels, illustrating the shift in the distribution of student performance between Cycle 1 and Cycle 2.

**Table 1. Comparison of Pre-test and Post-test Scores Across Cycles**

Cycle	Pre-test Mean	Pre-test Min/Max	Post-test Mean	Post-test Min/Max	Mastery (%)
Cycle 1	60	0 / 90	54	0 / 90	—
Cycle 2	50	0 / 85	64	0 / 95	76%

**Table 2. Classification of Post-test Performance Levels**

Classification	Score Range	Cycle 1 Post-test (%)	Cycle 2 Post-test (%)	Trend
Poor	0–60	35%	24%	↓
Fair	65–70	3%	0%	↓
Good	75–85	55%	52%	≈
Excellent	90–100	6%	24%	↑

### **Discussion**

The trajectory observed across the two cycles illustrates the value of iterative pedagogical refinement within action research, and offers a nuanced picture of how cartoon media's instructional effectiveness depends substantially on the broader pedagogical scaffolding within which it is deployed. The decline in Cycle 1 post-test scores, despite clearly increased behavioral engagement, suggests that affective engagement and motivation alone are insufficient to produce immediate gains in speaking performance when learners simultaneously lack the requisite vocabulary scaffolding to complete the assessed task. This finding aligns closely with Mayer's (2014) Cognitive Theory of Multimedia Learning, which cautions that multimedia input must be carefully managed to avoid overwhelming learners' processing capacity; absent pre-taught vocabulary, students in Cycle 1 likely faced excessive cognitive demands during the dubbing task, despite their evident enthusiasm for the cartoon content itself. This interpretation is further consistent with Thornbury's (2005) distinction between cognitive, affective, and performance-related dimensions of speaking difficulty, suggesting that Cycle 1's intervention successfully addressed the affective dimension of students' speaking difficulties while leaving the cognitive dimension, namely insufficient vocabulary knowledge, comparatively unaddressed.

The substantial improvement observed in Cycle 2, by contrast, corroborates Krashen's (1982) affective filter hypothesis particularly clearly. The introduction of structured vocabulary pre-teaching directly reduced the cognitive load associated with the dubbing task, while the dubbing activity itself continued to provide a low-stakes, enjoyable speaking context that appears to have further lowered students' affective filter, thereby facilitating more confident and fluent oral production. Student interview responses describing the Cycle 2 activities as enjoyable and confidence-building directly reflect this reduction in speaking anxiety, while the observed increase in participation, from 60% to 90%, provides corroborating behavioral evidence of the same underlying affective shift. The pronunciation and fluency gains observed in Cycle 2 further align with Paivio's (1986) Dual Coding Theory, as the dual encoding of spoken cartoon dialogue and corresponding animated visual content appears to have strengthened students' retention and subsequent recall of target vocabulary during speaking tasks, an effect that vocabulary pre-teaching may have further reinforced by establishing initial verbal encoding prior to the dual-coded multimedia exposure itself.

These findings are broadly consistent with prior research demonstrating the efficacy of cartoon and audio-visual media in improving speaking outcomes among young EFL learners, including studies documenting reduced speech errors and increased motivation through

cartoon and picture-cued media (Duran et al., 2021), measurable speaking gains through audio-visual cartoon applications among fifth-grade learners (Sandi et al., 2023), and improvements in vocabulary accuracy and pronunciation among beginner-level learners exposed to cartoon imagery (Karina et al., 2023). The present study extends this existing literature in a methodologically important respect: by employing a two-cycle action research design rather than a single-intervention study, it demonstrates that the pedagogical sequencing surrounding cartoon media, specifically the deliberate addition of vocabulary pre-teaching ahead of cartoon exposure and the structured use of dubbing as a practice modality, is critical to realizing cartoon media's full instructional potential. The use of cartoon media alone, as observed in Cycle 1, proved insufficient to produce measurable speaking gains, despite its evident success in fostering engagement and motivation; it was only once this affective engagement was paired with adequate cognitive scaffolding in Cycle 2 that the anticipated improvements in speaking performance materialized. This finding carries a clear practical implication for teachers: cartoon media should not be treated as a self-sufficient instructional intervention, but rather as one component within a deliberately sequenced lesson design that explicitly addresses learners' vocabulary needs in advance of, rather than solely during, multimedia exposure.

#### ***Pedagogical and Theoretical Implications***

At the theoretical level, the divergent outcomes of Cycle 1 and Cycle 2 offer a more granular test of the frameworks reviewed earlier than a single-intervention study would typically permit. Had this study employed only a single instructional cycle resembling Cycle 1, the resulting data might plausibly have been interpreted as partial disconfirmation of the affective filter hypothesis, given that increased engagement and reduced visible anxiety were not accompanied by improved post-test scores. The two-cycle design instead reveals a more precise picture: the affective filter hypothesis appears to hold most clearly once a baseline level of cognitive readiness, in this case adequate vocabulary knowledge, has been established, suggesting that affective and cognitive barriers to speaking operate as partially independent obstacles that may each require distinct, deliberately sequenced pedagogical responses rather than a single undifferentiated intervention.

At the pedagogical level, the findings carry direct implications for how cartoon media should be positioned within classroom instruction more broadly. Rather than treating cartoon viewing as the central instructional event, the present findings suggest that the activities surrounding cartoon exposure, namely vocabulary pre-teaching beforehand and structured, low-stakes production tasks such as dubbing afterward, may in fact be the more decisive determinants of whether measurable speaking gains are realized. This reframing is consistent with Abidin's (2015) three-stage pre-speaking, speaking, and post-speaking model, and suggests that teachers adopting cartoon media for speaking instruction should invest deliberate planning effort in the pre- and post-viewing stages rather than assuming that the cartoon content itself will carry the full instructional weight of the lesson.

At the institutional level, the relatively low cost and broad availability of cartoon media, requiring only a projector or display screen and freely or inexpensively accessible animated content, suggests that the barriers to implementing the refined instructional model demonstrated in this study are more a matter of pedagogical planning and teacher preparation than of financial or technological resourcing. This is a particularly salient

consideration for under-resourced Indonesian primary schools, where the present findings suggest that meaningful gains in speaking instruction may be achievable through more deliberate lesson sequencing alone, without necessarily requiring substantial new investment in technology or materials.

## CONCLUSION

This study demonstrates that cartoon media, when implemented through a structured, iterative Classroom Action Research design, can significantly improve the speaking skills of sixth-grade EFL learners in an Indonesian elementary school context. The trajectory of the two cycles offers an important methodological and pedagogical lesson: Cycle 1 revealed that increased engagement and motivation, while necessary, do not on their own guarantee immediate gains in measurable speaking performance, particularly when learners simultaneously face unaddressed vocabulary gaps. The targeted refinements introduced in Cycle 2, namely structured vocabulary pre-teaching and dubbing-based practice activities, directly addressed this gap, producing substantial improvements in both speaking performance and classroom participation that ultimately surpassed the study's predetermined mastery criterion of 75%.

These findings affirm the value of cartoon media as a low-cost, engaging, and pedagogically effective tool for developing oral communicative competence among young Indonesian EFL learners, while simultaneously underscoring the importance of deliberate instructional scaffolding in maximizing its effectiveness. The study's findings align with and extend established theoretical frameworks, including Krashen's (1982) affective filter hypothesis, Mayer's (2014) Cognitive Theory of Multimedia Learning, and Paivio's (1986) Dual Coding Theory, lending empirical support to the proposition that comprehensible, enjoyable, and well-scaffolded multimedia input creates the cognitive and affective conditions most conducive to speaking skill development among young learners.

On the basis of these findings, teachers seeking to implement cartoon-based speaking instruction are encouraged to incorporate structured vocabulary pre-teaching and dubbing activities as core, rather than incidental, components of their instructional design, and to consider pairing cartoon viewing with small-group or pair-based speaking activities to maximize equitable participation in larger classroom settings. It should be acknowledged that the present study is subject to certain limitations, including its restriction to a single classroom context, the relatively brief overall duration of the intervention, and the absence of a control group, each of which constrains the generalizability of the findings beyond the immediate research setting. Notwithstanding these limitations, the study makes a meaningful and original contribution to the growing body of literature on multimedia-assisted EFL speaking instruction in Indonesian primary education, providing both empirical evidence of cartoon media's pedagogical efficacy and a replicable, cycle-based instructional model that practitioners and future researchers may adapt and extend across comparable educational contexts.

Future research might productively extend these findings by examining the applicability of cartoon-based instruction across different elementary grade levels, by comparing the relative effectiveness of cartoon media against other multimedia modalities such as picture books or digital storytelling, and by investigating the long-term retention of speaking gains

through longitudinal follow-up assessment conducted several months beyond the conclusion of the intervention itself. Comparative experimental studies, in particular, would help disentangle the relative contribution of cartoon media itself from the pedagogical scaffolding, such as vocabulary pre-teaching, identified in this study as critical to its effectiveness. Such efforts would further strengthen the evidentiary basis for cartoon media's role within the broader landscape of multimedia-assisted EFL speaking instruction in Indonesian primary education.

## REFERENCES

- [1] Abidin, C. (2015). Communicative intimacies: Influencers and perceived interconnectedness. *Ada: A Journal of Gender, New Media, and Technology*, 8, 1–16.
- [2] Ariyanto, A., Priyayi, D. F., & Dewi, L. (2018). Penggunaan media pembelajaran biologi di Sekolah Menengah Atas (SMA) swasta Salatiga.
- [3] Binkley, M. R., Erstad, O., Herman, J. L., Raizen, S. A., Ripley, M., Miller-Ricci, M., & Rumble, M. (2012). Defining twenty-first century skills.
- [4] Bygate, M. (1987). *Speaking*. Oxford University Press.
- [5] Deterding, D., & Kirkpatrick, A. (2006). Emerging South-East Asian Englishes and intelligibility. *World Englishes*, 25(3-4), 391–409.
- [6] Duran, C., Aktay, E. G., & Kuru, O. (2021). Improving the speaking skill of primary school students instructed in a multigrade class through cartoons.
- [7] Gilakjani, A. P. (2012). The significant role of multimedia in motivating EFL learners' interest in English language learning. *International Journal of Modern Education and Computer Science*, 4, 57–66.
- [8] Karina, S. S. I., Wuriyanto, A. B., & Prihatini, A. (2023). Penerapan media gambar kartun pada pembelajaran keterampilan berbicara BIPA tingkat A1. *Stilistika: Jurnal Pendidikan Bahasa dan Sastra*.
- [9] Kemmis, S., & McTaggart, R. (1988). *The action research planner*. Deakin University Press.
- [10] Krashen, S. D. (1982). *Principles and practice in second language acquisition*. Pergamon Press.
- [11] Lisnawati, I., Yuniawati, Y., & Kusmini, T. (2017). Student presentations upgrades through use of language learning strategies.
- [12] Mayer, R. E. (2014). *The Cambridge handbook of multimedia learning*. Cambridge University Press.
- [13] Paivio, A. (1986). *Mental representations: A dual coding approach*. Oxford University Press.
- [14] Sandi, B. A., Yuhana, Y., & Andriana, E. (2023). Pengembangan media audio visual melalui aplikasi Cartoon Story Maker untuk meningkatkan keterampilan berbicara di kelas V Sekolah Dasar. *Jurnal Pendidikan Dasar Perkhasa*.
- [15] Thornbury, S. (2005). *How to teach speaking*. Pearson Education.
- [16] Yuniwati, R., Wijaya, B. S., & Rosnija, E. (2015). Improving students' speaking ability through short drama.
- [17] Zhen, Z. (2016). The use of multimedia in English teaching. *US-China Foreign Language*, 14, 182–189.

- [18] Young, D. J. (1991). Creating a low-anxiety classroom environment: What does language anxiety research suggest? *The Modern Language Journal*, 75(4), 426–439.

HALAMAN INI SENGAJA DIKOSONGKAN