Arus Jurnal
Psikologi dan Pendidikan (AJPP)
Website: http://jurnal.ardenjaya.com/index.php/ajpp
Email: jurnal.ajpp@ardenjaya.com

The Effect of Movie Scene Clips on Seventh-Grade Students' English-Speaking Performance: A Quasi-Experimental Study in an Indonesian EFL Context

INFO PENULIS

M. Reza Pahlevi
Universitas Negeri Surabaya
reza.22192@mhs.unesa.ac.id

Sueb
Universitas Negeri Surabaya
sueb@unesa.ac.id

Fulviana Arraisah
Universitas Negeri Surabaya
its.me.raisaa@gmail.com

INFO ARTIKEL

ISSN: 2963-8933
Vol. 5, No. 2 Juni 2026
http://jurnal.ardenjaya.com/index.php/ajpp

© 2026 Arden Jaya Publisher All rights reserved

Saran Penilisan Referensi:

Pahlevi, M. R., Sueb, Arraisah, F. (2026). The Effect of Movie Scene Clips on Seventh-Grade Students' English-Speaking Performance: A Quasi-Experimental Study in an Indonesian EFL Context. *Arus Jurnal Psikologi dan Pendidikan, 5* (2), 2546-2555.

Abstrak

Penelitian ini bertujuan untuk mengetahui komponen performa berbicara yang memperoleh peningkatan N-Gain tertinggi pada siswa yang diajar menggunakan potongan adegan film (movie scene clips) serta untuk menguji apakah terdapat perbedaan peningkatan yang signifikan antara siswa yang diajar menggunakan potongan adegan film dan siswa yang diajar menggunakan aktivitas berbicara berbasis buku teks. Penelitian ini menggunakan desain kuasi-eksperimen dengan pola non-equivalent control group yang melibatkan 42 siswa kelas tujuh pada sebuah sekolah menengah pertama di Surabaya, terdiri atas kelas eksperimen dan kelas kontrol. Kelas eksperimen memperoleh perlakuan berupa potongan adegan film Zootopia berdurasi tiga sampai empat menit melalui teknik menyimak tiga tahap, sedangkan kelas kontrol mengikuti pembelajaran berbicara konvensional berbasis buku teks dengan topik yang sama. Performa berbicara diukur melalui tes lisan sebelum dan sesudah perlakuan yang dinilai oleh dua penilai menggunakan rubrik analitis adaptasi CEFR level A2 pada lima komponen, yaitu akurasi pelafalan, kelancaran, intonasi, kepercayaan diri, dan ekspresi. Data dianalisis dengan skor Normalized Gain (N-Gain), uji normalitas Shapiro-Wilk, dan uji Mann-Whitney U. Hasil penelitian menunjukkan bahwa akurasi pelafalan, kelancaran, dan kepercayaan diri memperoleh rerata N-Gain tertinggi dan identik sebesar 0,92, diikuti intonasi (0,80) dan ekspresi (0,77); seluruhnya berada pada kategori efektivitas tinggi. Uji Mann-Whitney U menunjukkan perbedaan yang signifikan antara kedua kelompok pada seluruh komponen ($p = 0{,}000 < 0{,}05$), sementara kelas kontrol hanya berada pada kategori rendah. Temuan ini menunjukkan bahwa potongan adegan film dapat menjadi media pembelajaran yang efektif untuk mengembangkan aspek linguistik sekaligus aspek afektif keterampilan berbicara siswa sekolah menengah pertama.

Kata Kunci: Potongan Adegan Film, Performa Berbicara, Kuasi-Eksperimen, Skor N-Gain, Siswa Sekolah Menengah Pertama, Pembelajaran Bahasa Inggris Sebagai Bahasa Asing.

Abstract

This study aims to identify the speaking performance component with the highest N-Gain among students taught using movie scene clips and to examine whether there is a significant difference in N-Gain scores between students taught using movie scene clips and those taught using textbook-based speaking activities. A quasi-experimental non-equivalent control group design was employed, involving 42 seventh-grade students of a junior high school in Surabaya who were assigned into an experimental class and a control class. The experimental class received a treatment using three-to-four-minute movie scene clips from Zootopia delivered through a three-phase viewing technique, while the control class followed textbook-based speaking instruction on the same topic. Speaking performance was measured through oral tests scored by two independent raters using an analytic rubric adapted from the CEFR A2 descriptors covering pronunciation accuracy, fluency, intonation, self-confidence, and expression. The data were analysed using Normalized Gain (N-Gain) scores, the Shapiro-Wilk normality test, and the Mann-Whitney U test. The findings revealed that pronunciation accuracy, fluency, and self-confidence obtained the highest and identical mean N-Gain of 0.92, followed by intonation (0.80) and expression (0.77), all of which fall into the high effectiveness category. The Mann-Whitney U test indicated a significant difference between the two groups across all five components ($p = 0.000 < 0.05$), whereas the control class remained within the low effectiveness category. These findings suggest that movie scene clips can serve as an effective instructional medium for developing both the linguistic and the affective dimensions of junior high school students' speaking performance.

Keywords: Movie Scene Clips, Speaking Performance, Quasi-Experimental Study, N-Gain Score, Junior High School Students, English As A Foreign Language.

A. Introduction

Speaking has long been regarded as the most demanding skill for learners of English as a foreign language, since it requires learners to process meaning and produce it orally in real time. In the era of global communication, the ability to speak clearly and confidently is no longer an optional competence but a practical necessity. Within the boundaries of the EFL classroom, however, speaking constitutes a real and intimidating hurdle for both the teacher and the learner (Bouzar, 2019). Teachers are required to find appropriate and innovative procedures to scaffold learners, while learners are expected to find engaging ways to practise the target language.

In the Indonesian context, particularly at the junior high school level, speaking often receives less attention than reading and writing. Oral communication remains a challenging area because of limited classroom time and a teaching approach that largely remains teacher centred, which results in very limited opportunities for students to engage in meaningful speaking practice (Setiawan et al., 2025). As a consequence, many students use English almost exclusively inside the classroom and rarely experience situations in which English functions as a genuine means of communication.

The obstacles faced by junior high school students are both linguistic and psychological. Linguistically, learners struggle with limited vocabulary, inaccurate pronunciation, and inadequate control of grammar; psychologically, they experience anxiety, hesitation, and low self-confidence (Jamoom & Bahron, 2024; Rofi'i, 2023). Classroom realities in many Indonesian secondary schools show that a large proportion of students possess low baseline capabilities in oral production and consistently struggle with accurate pronunciation, structural grammar, and contextual vocabulary selection because of internal and environmental constraints (Hotmaria, 2021). Anxiety related to pronunciation in particular has been reported to restrict learners' willingness to speak and to reduce the quality of their oral output (Yousif, 2025), while affective barriers such as fear of making mistakes make many learners perceive speaking as the most difficult English skill (Yuliana & Puspita, 2021).

These conditions underline the need for instructional media that can simultaneously provide authentic linguistic models and lower the affective barriers that inhibit oral production. Among the available multimedia resources, movie scene clips, that is short excerpts of English-language films, are particularly promising because they display real language use together with

natural intonation, facial expressions, and gestures (Hoinbala, 2022). Their brevity allows teachers to focus on specific language features within a limited allocation of time while maintaining an enjoyable classroom atmosphere that motivates students to imitate native-like delivery (Siahaan et al., 2024). Through film clips, learners hear English as it is actually used in situated interaction, which supports conversational skills and the comprehension of spoken language (Rahmah & Pranata, 2024; Kholmurodovna, 2024). English films have also been reported to improve practical conversation skills, acoustic understanding, and authentic pronunciation models, and to be perceived by learners as more motivating than conventional textbook drills (Shaheen et al., 2026).

The use of movie scene clips in the present study is grounded in four interrelated theoretical frameworks. The Cognitive Theory of Multimedia Learning explains that learners gain more from words combined with pictures than from words alone, because information is processed through two separate channels, the auditory or verbal channel and the visual or pictorial channel. Movie scene clips activate both channels simultaneously through spoken dialogue, intonation, and background sound on the one hand, and action, facial expression, and subtitles on the other, thereby supporting retention of spoken language patterns. Krashen's Input Hypothesis complements this account by proposing that acquisition occurs when learners receive comprehensible input slightly beyond their current level; authentic film dialogue offers such input while its low-anxiety and enjoyable nature lowers the affective filter. Communicative Language Teaching supplies the productive stage of the process, since the clips are not merely watched but used as stimuli for output through role play and scene reenactment. Finally, Long's Interaction Hypothesis justifies the use of paired performance tasks, as acquisition is activated when learners negotiate meaning, check comprehension, and modify their output during peer interaction. Recent evidence on multisensory instruction further supports the claim that integrating auditory and visual cues facilitates the acquisition of second-language phonological features (Li et al., 2026).

A number of studies have examined the contribution of audiovisual media to speaking instruction. Hariati (2022) reported that English films improved students' speaking skills, while Hasyim and Tarihoran (2024) found that YouTube materials affected speaking performance positively. Junita and Prasad (2024) and Fadilla and Mutiarani (2024) demonstrated that animation videos and animated series supported the development of speaking ability and fluency, and Hidayah (2022) documented the usefulness of YouTube movie clips for comprehension of spoken input. Most of these studies, however, either measured speaking as a single global construct, employed full-length films, or relied on qualitative and classroom action research designs. Consequently, relatively little quantitative evidence is available on how short movie scene clips affect specific micro-components of speaking performance among young adolescent learners.

A further methodological gap concerns the way improvement is measured. In school-based research, intact classes rarely start from equivalent baselines, yet many studies still compare raw post-test scores, which risks attributing pre-existing differences to the treatment. Analysing Normalized Gain (N-Gain) scores rather than post-test means offers a fairer estimate of actual improvement under such conditions and has not been widely applied in studies of film-based speaking instruction in Indonesian junior high schools.

The present study therefore investigates the effect of movie scene clips on five components of English-speaking performance, namely pronunciation accuracy, speaking fluency, intonation, self-confidence, and expression, among seventh-grade students in an Indonesian EFL context. Two research questions guide the study. First, which speaking performance component shows the highest N-Gain among students taught using movie scene clips? Second, is there a significant difference in N-Gain scores between students taught using movie scene clips and those taught using textbook-based speaking activities?

B. Methodology

Research Design

This study employed a quantitative approach with a quasi-experimental design to examine the causal relationship between the independent variable, the use of movie scene clips, and the dependent variable, students' English-speaking performance. A non-equivalent control group design was applied because intact classes of seventh-grade students were assigned to either the experimental or the control group without disrupting the institutional school schedule. Both groups took a pre-test before the treatment period and a post-test after the instructional

sessions, so that the improvement of each group could be compared under natural classroom conditions in which full randomisation is not feasible.

A central methodological principle in this design was the maintenance of a uniform construct across different administration formats. Both the pre-test and the post-test measured identical linguistic and psychological dimensions, namely pronunciation accuracy, speaking fluency, intonation, self-confidence, and expression. The difference between the two testing occasions concerned only the mode of administration and not the constructs assessed, the task architecture, or the scoring rubric.

Participants

The participants were seventh-grade students of a junior high school in Surabaya during the 2025/2026 academic year. Two intact classes were selected through purposive sampling based on the recommendation of the English teacher and the availability of comparable classroom schedules, with the consideration that both classes had relatively similar initial levels of English proficiency. One class served as the experimental group, which was taught using movie scene clips, and the other served as the control group, which was taught using conventional textbook-based speaking activities. The study initially involved 50 students; however, because several students were absent during the data collection process, only 42 students completed all research procedures and were included in the final analysis. All participants were EFL learners who used Indonesian or a local language for daily communication outside the English classroom.

Instruments

The main instrument of this study was a speaking test administered before and after the treatment. The test was designed by the researcher using speaking performance indicators relevant to junior high school EFL learners and was validated by English teachers through an expert judgement checklist. The test consisted of three operational sections. In the first section, students read aloud a structured text prepared by the researcher, a task designed to evaluate the linguistic mechanics of speech production, particularly pronunciation and intonation. In the second section, students retold the story in their own words, which measured speaking fluency in terms of natural flow and reduced hesitation as well as expression in terms of emotional and contextual delivery. In the third section, students responded to two open-ended opinion questions related to the topic of the text, which focused on self-confidence in spontaneous communication.

Students' oral performance was assessed using an analytic rubric adapted from the CEFR descriptors for the A2 level. The core criteria of phonological control and spoken fluency were adopted from the CEFR, while three further criteria were added in order to capture the multidimensional effect of audiovisual learning. The final rubric therefore contained five indicators. Pronunciation accuracy measured the accurate production of sounds and word stress. Speaking fluency measured speech flow, pacing, and the management of natural pauses. Intonation measured voice modulation, pitch control, and rhythm in the expression of meaning. Self-confidence, added by the researcher, measured psychological readiness to speak without disruptive performance anxiety. Expression, also added by the researcher, measured the learners' use of emotional delivery, facial gestures, and contextual cues modelled in the clips. Each indicator was scored on a four-point scale ranging from one (poor) to four (excellent). Two qualified raters scored every performance independently, and the degree of agreement between raters was examined using Pearson Product-Moment Correlation, with a coefficient of 0.70 or higher taken as the threshold of reliability.

Data Collection Procedure

Data collection was conducted systematically and sequentially in three steps. In the first step, both groups took an individual oral pre-test measuring their initial performance across the five speaking components. The researcher distributed a prepared reading text to all students and allowed an identical preparation time, after which students were called forward one by one according to the attendance list to perform three sequential tasks: reading the text aloud, retelling the story without looking at the text, and answering two open-ended opinion questions posed by the researcher.

The second step consisted of the pedagogical treatment, which spanned two weeks with two sessions per week. The experimental group was taught using a short animated movie scene clip of three to four minutes taken from Zootopia, specifically the first meeting between Judy Hopps and Nick Wilde, which was selected because of its age-appropriate theme, standard vocabulary, natural speech pace, high visual clarity, and clear dialogue at a level suitable for CEFR A2 learners. Each treatment session followed a three-phase viewing technique. The pre-

viewing phase activated background knowledge; the while-viewing phase involved repeated playbacks and intentional pauses during which students performed shadowing and choral repetition in order to imitate native-like pronunciation and to analyse the rhythm and pacing required for fluency; and the post-viewing phase required students to work in pairs and to perform character-based scene reenactments in which they embodied one of the characters and imitated the delivery of the original speakers. The control group, in contrast, engaged in standard speaking exercises drawn strictly from the existing school textbook under a conventional teacher-centred format, focusing on the same overarching topic of fable texts in order to maintain thematic consistency. No performance data were collected during this phase.

In the third step, both groups took a paired oral post-test. Students were grouped into pairs and received a new reading text that introduced a different story and different character dialogues, calibrated to the same CEFR A2 level as the pre-test. Each pair performed three sequential tasks: reading the dialogue aloud in alternating turns, retelling the story collaboratively without the script, and responding alternately to two open-ended opinion questions. Although the tasks were performed in pairs, the analytic scoring remained individual: the unit of analysis was the individual learner, and each student was evaluated independently against the same rubric and the same performance indicators used in the pre-test. Both testing sessions were scored on site by two independent raters, which allowed immediate assessment without recourse to external audio or video recordings.

Methodological Adjustments and Field Constraints

Because the study was conducted in a natural school setting, two constraints required pragmatic adjustment and are reported here in the interest of transparency. The first concerns the administration format of the post-test. The pre-test was administered individually, but during the intervention phase the allocated English instruction hours in the experimental class were unexpectedly reallocated to school-wide events for two consecutive days. Given the reduced instructional time and the rigidity of the academic calendar, administering an individual post-test for all students was no longer administratively feasible, and the post-test was therefore delivered as a paired performance task. This format is defensible on three grounds. Modern language assessment recognises that speaking competence is not merely monologic but includes interactional competence, namely the ability to co-construct meaning and manage turns with an interlocutor. Speaking in pairs also functions as a psychological buffer that may reduce language anxiety among young adolescent learners who are often intimidated by individual performance. Furthermore, communicative tasks correspond to the interactive nature of the movie scene clips used in the treatment. It must nevertheless be acknowledged that interactional support from a peer may inflate fluency and self-confidence scores, and this possibility is treated as a limitation of the study.

The second constraint concerns baseline equivalence. The pre-test data showed that the scores of the two groups were not entirely equal, with the experimental group starting from a higher position. This initial difference is a known drawback of non-randomised designs that rely on intact classes. In order to avoid attributing pre-existing advantages to the treatment, the analysis was conducted on Normalized Gain scores, which express the improvement achieved relative to the room for improvement available to each learner, rather than on raw post-test means.

Data Analysis Technique

The data were analysed quantitatively with the assistance of SPSS through four stages. First, inter-rater reliability was examined using Pearson Product-Moment Correlation in order to ensure that the two raters scored consistently. Second, descriptive statistics, including means and percentages, were calculated for the overall speaking performance and for each of the five components. Third, the Normalized Gain score was computed for every student on each component. The gain value ranges from -1.00 to 1.00 and is classified into three effectiveness categories, namely high for g greater than or equal to 0.70, medium for g between 0.30 and 0.70, and low for g below 0.30. Fourth, the distribution of the N-Gain scores was examined using the Shapiro-Wilk test, which was selected because the number of participants in each class was fewer than fifty, and the hypothesis was then tested using the appropriate comparison procedure. A percentage growth formula was additionally used to describe the absolute increase of each component from the pre-test to the post-test.

C. Findings and Discussion

Findings

This section presents the results of the data analysis obtained from the experimental class and the control class. In order to maintain consistency of statistical logic and to answer the research questions, the inferential analysis was conducted directly on the Normalized Gain scores rather than on the post-test scores, so that the evaluation reflects the actual improvement of each learner while taking the difference in starting levels into account.

Normality of the N-Gain Scores

The distribution of the N-Gain scores for the five speaking components was examined using the Shapiro-Wilk test at a significance level of 0.05. The results showed that nearly all components in both classes produced p-values below this threshold. In the experimental class, the significance values for pronunciation accuracy, speaking fluency, self-confidence, and intonation were all 0.000, indicating a clear violation of the normality assumption, while the significance value for expression was 0.006, which also falls below 0.05. Although the N-Gain distribution of expression in the control class was normal, its counterpart in the experimental class was not. Since the assumption of normality was violated in at least one of the groups for every component, a parametric comparison could not be justified, and the non-parametric Mann-Whitney U test was therefore applied to the N-Gain scores.

Differences in N-Gain Scores between the Two Groups

Table 1 presents the results of the Mann-Whitney U test computed on the N-Gain scores of the five speaking components. The test was conducted to evaluate whether the implementation of movie scene clips in the experimental class produced a statistically significant difference in learning improvement compared with the textbook-based instruction used in the control class. The null hypothesis is rejected when the asymptotic significance value is lower than the alpha level of 0.05.

Table 1. Mann-Whitney U Test on the N-Gain Scores

Test Statistics	N-Gain Pronunciation	N-Gain Fluency	N-Gain Confidence	N-Gain Expression	N-Gain Intonation
Mann-Whitney U	0.000	2.000	3.000	11.500	15.000
Wilcoxon W	300.000	302.000	303.000	311.500	315.000
Z	-6.171	-5.981	-5.917	-5.829	-5.519
Asymp. Sig. (2-tailed)	0.000	0.000	0.000	0.000	0.000

As shown in Table 1, the asymptotic significance values for all five components are 0.000, which is considerably lower than the alpha level of 0.05. The null hypothesis is therefore rejected and the alternative hypothesis is accepted for every component. The Z values, which range from -5.519 for intonation to -6.171 for pronunciation accuracy, indicate a consistent and substantial separation between the two distributions. The results demonstrate that the difference between the two groups lies not merely in their final scores but in their rate of improvement, which confirms that the treatment produced a significantly better progression in oral proficiency across both the linguistic and the affective sub-skills.

Mean N-Gain of Each Speaking Component

Table 2 presents the mean Normalized Gain of the experimental and control classes for each of the five speaking components together with the corresponding effectiveness category.

Table 2. Mean N-Gain of Each Speaking Component

No	Speaking Component	Mean N-Gain (Experimental)	Mean N-Gain (Control)	Effectiveness Category
1	Pronunciation accuracy	0.92	0.07	Experimental: high; control: low
2	Speaking fluency	0.92	-0.02	Experimental: high; control: low
3	Intonation	0.80	0.04	Experimental: high; control: low
4	Self-confidence	0.92	0.12	Experimental:

5	Expression	0.77	-0.02	high; control: low Experimental: high; control: low

The improvement achieved by the experimental class placed all five components in the high effectiveness category. Pronunciation accuracy, speaking fluency, and self-confidence obtained an identical mean N-Gain of 0.92, which represents the highest gain recorded in this study, while intonation and expression followed closely with mean gains of 0.80 and 0.77 respectively. These figures indicate that the intervention supported the development of both the linguistic and the psychological dimensions of speaking simultaneously. The control class, in contrast, remained within the low effectiveness category on all indicators, with mean gains of 0.12 for self-confidence, 0.07 for pronunciation accuracy, and 0.04 for intonation. Speaking fluency and expression in the control class even produced negative gain values of -0.02, which reflects a slight decrease in the post-test results.

Percentage Increase from Pre-Test to Post-Test

Table 3 displays the comparative percentage means of students' speaking performance in the pre-test and the post-test sessions, together with the real increase expressed in percentage points.

Table 3. Pre-Test and Post-Test Percentage Means and Real Increase

No	Speaking Component	Class	Pre-test (%)	Post-test (%)	Real increase (% points)
1	Pronunciation accuracy	Experimental	81.00	98.50	+17.50
		Control	57.29	60.42	+3.13
2	Speaking fluency	Experimental	81.00	98.50	+17.50
		Control	53.65	52.60	-1.05
3	Self-confidence	Experimental	80.50	98.50	+18.00
		Control	56.77	61.79	+5.02
4	Intonation	Experimental	77.50	95.50	+18.00
		Control	34.38	36.98	+2.60
5	Expression	Experimental	48.00	88.00	+40.00
		Control	36.98	35.94	-1.04

The descriptive breakdown confirms the pattern identified in the inferential analysis. In the experimental class, the largest absolute growth occurred in expression, which rose from 48.00 per cent to 88.00 per cent, an increase of 40.00 percentage points. Pronunciation accuracy and speaking fluency both advanced from 81.00 per cent to 98.50 per cent, an increase of 17.50 percentage points, while self-confidence rose from 80.50 per cent to 98.50 per cent and intonation from 77.50 per cent to 95.50 per cent, each gaining 18.00 percentage points. It is worth noting that the components with the highest mean N-Gain are not identical with those showing the largest absolute growth: because pronunciation, fluency, and self-confidence already started from a comparatively high baseline, their smaller absolute increase still represents the near-complete closure of the distance to the maximum score, which is precisely what the Normalized Gain is designed to capture.

In the control class the growth was limited. The increases remained below ten percentage points, amounting to 5.02 points for self-confidence, 3.13 points for pronunciation accuracy, and 2.60 points for intonation. Two components even declined slightly, with speaking fluency falling from 53.65 per cent to 52.60 per cent and expression from 36.98 per cent to 35.94 per cent. This descriptive pattern indicates that textbook-based activities, as implemented in this study, contributed considerably less to the improvement of oral performance than the clip-based activities did.

Discussion

The findings of this study corroborate and extend a broad range of previous research on EFL speaking instruction. The significant gains recorded in the experimental group are consistent with the claim that authentic multimedia supports the acquisition of suprasegmental features such as natural intonation that are rarely acquired through paper-based drills, and with the argument advanced by Hoinbala (2022) that authentic video input creates a natural communicative context for oral production. They also reinforce the results reported by Hariati

(2022), Rahmah and Pranata (2024), and Shaheen et al. (2026), who found that film media positively influenced pronunciation, intonation, and self-confidence, as well as the findings of Junita and Prasad (2024) and Fadilla and Mutiarani (2024) concerning the contribution of animated audiovisual materials to speaking ability and fluency.

At the same time, the present study diverges from earlier work in two important respects. Whereas most previous studies measured macro-level oral proficiency, this study isolates and quantifies the impact of short clips on five micro-dimensions of speaking performance among seventh-grade learners. Furthermore, while a considerable proportion of the existing literature relies on qualitative or classroom action research designs, or employs full-length films, the present study demonstrates through a quasi-experimental design that clips of only three to four minutes are sufficient to produce measurable gains. The brevity of the clip appears to be an advantage rather than a limitation, since it provides a dense multi-sensory environment while keeping the cognitive load manageable through contextualised lexical chunks.

The identical mean N-Gain of 0.92 obtained for pronunciation accuracy, speaking fluency, and self-confidence deserves particular attention because it indicates that the linguistic and the affective dimensions of speaking improved in parallel rather than sequentially. This pattern is consistent with the theoretical framework adopted in this study. The shadowing and choral repetition activities in the while-viewing phase provided repeated exposure to phonological models, which aligns with the multisensory account of second-language phonological learning (Li et al., 2026), while the enjoyable and low-pressure atmosphere created by the animated material plausibly lowered the affective filter and reduced the pronunciation anxiety documented by Yousif (2025). The role-play tasks in the post-viewing phase then converted this input into output under conditions of peer support, which is consistent with the view that speaking fluency develops most effectively through interactive and task-based practice (Masuram & Sripada, 2020; Setiawan et al., 2025).

The comparatively lower, though still high, gains for intonation and expression can be interpreted in a similar manner. Intonation involves the control of pitch and rhythm and is acknowledged in the literature as a feature that develops gradually and requires sustained exposure (Adawiyah et al., 2023). Expression, in turn, showed the largest absolute increase in the experimental class, which suggests that the visual modelling of facial gestures and emotional delivery in the clip addressed an aspect of oral performance that textbook instruction leaves almost entirely untouched. The negative gains for fluency and expression in the control class support this interpretation: without dynamic audiovisual modelling, learners had no reference for expressive delivery and their performance in these components did not progress.

Several limitations should be considered when interpreting these results. First, the two classes did not start from equivalent baselines, and although the use of Normalized Gain scores mitigates this problem, it does not eliminate the possibility that unmeasured class-level differences contributed to the outcome. Second, the change in the administration format of the post-test means that peer support may have inflated the scores for fluency and self-confidence, even though scoring remained individual and the rubric was held constant. Third, the treatment lasted only two weeks and involved a single film clip, so the durability of the observed gains cannot be established from these data. Fourth, the participants were drawn from two intact classes in a single school, which limits the generalisability of the findings. Future studies would benefit from longer treatment periods, a wider variety of clips, delayed post-tests, and where possible a randomised or matched design.

D. Conclusion

This study examined the effect of movie scene clips on the English-speaking performance of seventh-grade students in an Indonesian EFL context. With respect to the first research question, the descriptive analysis showed that pronunciation accuracy, speaking fluency, and self-confidence shared the highest mean Normalized Gain of 0.92, followed by intonation with 0.80 and expression with 0.77. All five components therefore fell within the high effectiveness category, which indicates that the systematic use of movie scene clips supports the linguistic and the psychological dimensions of oral performance simultaneously.

With respect to the second research question, the Mann-Whitney U test conducted on the N-Gain scores revealed a significant difference between the two groups across all five components, with asymptotic significance values of 0.000, which is well below the alpha level of 0.05. While the experimental class reached the high effectiveness range on every sub-skill, the control class remained within the low effectiveness range, with gains of 0.12 for self-confidence,

0.07 for pronunciation accuracy, and 0.04 for intonation, and with slight decreases of -0.02 for both fluency and expression.

These results carry several implications. Theoretically, they support the principles of comprehensible input and multi-sensory learning in second-language acquisition, since the simultaneous provision of auditory and visual stimuli appears to facilitate the retention of language chunks and the imitation of natural speech rhythm during active performance. Methodologically, the study shows that the underlying speaking constructs can be assessed consistently even when the administration format has to change from individual to paired performance, provided that the analytic rubric, the rater standards, and the evaluative criteria remain strictly identical. Practically, the findings suggest that English teachers may integrate short film clips as an accessible alternative to text-heavy instruction, that students may use audiovisual media independently outside the classroom to sustain their practice, and that researchers working in active school settings should allow additional flexibility in their timelines and rely on Normalized Gain rather than raw score comparison when intact classes present unequal baseline conditions.

E. References

Adawiyah, R., Barus, S. K., & Damanik, E. S. D. (2023). An analysis student's intonation in reading English newspaper. Jurnal Pendidikan Bahasa Inggris Undiksha, 11(3), 304-309. https://doi.org/10.23887/jpbi.v11i3.43141

Arifin, Mafulah, S., & Sholeh, A. (2025). EFL learner's strategies to build self-confidence in speaking: A narrative inquiry. Trans-Kata: Journal of Language, Literature, Culture, and Education, 5(2), 461-471. http://ejournal.transbahasa.co.id/index.php/jllce

Bouzar, S. (2019). Issues in teaching speaking to EFL learners. Education and Linguistics Research, 5(1), 68-79. https://doi.org/10.5296/elr.v5i1.14705

Edmondson, W. J., House, J., & Kadar, D. Z. (2023). Expressions, speech acts and discourse: A pedagogic interactional grammar of English. Cambridge University Press.

Fadilla, D. I., & Mutiarani, M. (2024). Animating English fluency: Empowering senior high students through speaking activities with animated series. Jurnal Studi Guru dan Pembelajaran, 7(1), 419-428. https://doi.org/10.30605/jsgp.7.1.2024.3564

Goh, C. C., & Burns, A. (2012). Teaching speaking: A holistic approach. Cambridge University Press.

Hariati, P. (2022). Improving students' speaking skill through English movie. Language Literacy: Journal of Linguistics, Literature, and Language Teaching, 6(2), 454-461. https://doi.org/10.30743/ll.v6i2.6176

Hasyim, N. U., & Tarihoran, N. (2024). The widely common media to teach English: Does YouTube affect students' speaking skills? Indonesian Journal of English Language Studies (IJELS), 10(2), 118-126. https://doi.org/10.24071/ijels.v10i2.9049

Hidayah, N. (2022). The use of YouTube movie clips to improve students' listening comprehension. Journal of English Language Teaching and Linguistics, 7(2), 133-145. https://doi.org/10.52166/edulitics.v7i1.3028

Hoinbala, F. R. (2022). Movies as an authentic input in L2 speaking class: A dynamic usage-based approach in EFL teaching in Indonesia. International Journal of Language Education, 6(1), 1-9. https://doi.org/10.26858/ijole.v6i1.20196

Hotmaria, H. (2021). Upaya meningkatkan keterampilan berbicara bahasa Inggris pada materi pengandaian diikuti perintah/saran menggunakan strategi pembelajaran Three Step Interview. Journal of Education Action Research, 5(1), 1-9. https://doi.org/10.23887/jear.v5i1.31215

Jamoom, O. A., & Bahron, N. O. (2024). From hesitation to fluency: Unraveling speaking barriers in EFL context. International Journal of English Literature and Social Sciences, 9(5), 721-737. https://doi.org/10.22161/ijels.9.5

Junita, D. N., & Prasad, R. D. (2024). The effect of using animation videos on students' speaking ability. Journal of Language, Literature, and Educational Research, 1(2), 39-44. https://doi.org/10.37251/jolle.v1i2.1063

Kholmurodovna, D. S. (2024). Speaking fluency development in English as a Foreign Language (EFL) contexts. World Bulletin of Social Sciences (WBSS), 32(March), 1-5. https://www.scholarexpress.net

Li, J., Zhang, Q., & Wang, Z. (2026). Multisensory teaching for second-language acquisition of voicing stop consonants: Integrating auditory, visual, and tactile cues. Language Teaching Research. https://doi.org/10.1177/13621688261443729

Ma, Q., Mei, F., & Qian, B. (2026). Exploring EFL students' pronunciation learning supported by corpus-based language pedagogy. Computer Assisted Language Learning, 39(4), 915-941. https://doi.org/10.1080/09588221.2024.2432965

Masuram, J., & Sripada, P. N. (2020). Developing spoken fluency through task-based teaching. Procedia Computer Science, 172, 623-630. https://doi.org/10.1016/j.procs.2020.05.080

Nurhayati, D. A. W., Djatmika, Santosa, R., & Wiratno, T. (2016). Indonesian influence in developing speaking skill in learning English: EFL learners' impediments. Proceedings of International Seminar PRASASTI III: Current Research in Linguistics, 67-96. https://doi.org/10.20961/pras.v0i0.1488

Premarathne, P. B. T. K., & Nagodavithana, S. (2024). Developing oral fluency of second language learners: A narrative review. ISRG Journal of Arts, Humanities and Social Sciences, 2(4), 19-28. https://isrgpublishers.com/isrgjahss

Rahmah, L. N. A., & Pranata, M. S. A. (2024). Students' experiences in using English movies to improve students' speaking skills. Journal of English Language and Education, 9(4), 142-153. https://doi.org/10.31004/jele.v9i4.527

Rita, D. (2022). Peningkatan keterampilan berbahasa Inggris menggunakan Gallery Exhibition Project dalam pembelajaran bahasa Inggris pada topik report text. Journal of Educational Learning and Innovation (ELIa), 2(1), 96-120. https://doi.org/10.46229/elia.v2i1

Rodriguez-Fuentes, R. A., & Calle-Diaz, L. (2023). Fluency development beyond speech rate: A study on the effects of read aloud protocols on EFL learners. Ampersand, 11, 100133. https://doi.org/10.1016/j.amper.2023.100133

Rofi'i, A. (2023). Kesulitan berbicara siswa sekolah dasar dalam pembelajaran bahasa Inggris. Jurnal Elementaria Edukasi, 6(4), 1895-1904. https://doi.org/10.31949/jee.v6i4.6851

Santiana, S., Yulianti, N., Syakira, S., Khan, T., & Siripipathanakul, S. (2026). Language learning strategies for managing speaking anxiety among English students: A narrative inquiry. PANYONARA: Journal of English Education, 8(1), 70-88.

Setiawan, A., Rahman, A., Jupri, & Meraj, M. (2025). Exploring the role of peer feedback in enhancing EFL learners' English speaking fluency and accuracy. JOLLS: Journal of Language and Literature Studies, 5(2), 461-471. https://doi.org/10.36312/jolls.v5i2.2868

Shaheen, R., Tumrani, A., & Maitlo, S. K. (2026). The impact of English movies on speaking fluency: A study of college-level ESL learners. The Critical Review of Social Sciences Studies, 4(1), 742-754. https://doi.org/10.59075/z04dr552

Siahaan, S., Eriyanti, R. W., Rofieq, A., & Huda, A. M. (2024). Students' perception regarding the integration of English-subtitled films to enhance English language learning. Jurnal Cahaya Pendidikan, 10(1), 67-76. https://doi.org/10.33373/chypen.v10i1.5948

Vy, L. T. M., Huong, T. L. T., Quy, T. N., Cuong, V. Q., & Anh, N. T. (2024). Determined factors and effective strategies for developing English speaking fluency among Vietnamese university students. PASAA Journal, 68, 32-63.

Wang, W., Rezaei, Y. M., & Izadpanah, S. (2024). Speaking accuracy and fluency among EFL learners: The role of creative thinking, emotional intelligence, and academic enthusiasm. Heliyon, 10, e37620. https://doi.org/10.1016/j.heliyon.2024.e37620

Yousif, A. S. A. (2025). Pronunciation anxiety among EFL learners: Causes, consequences, and classroom interventions. Journal of Language Teaching and Research, 16(6), 1955-1963. https://doi.org/10.17507/jltr.1606.17

Yuliana, R., & Puspita, D. (2021). The psychological filter: Why speaking English is a nightmare for young learners. Journal of Affective Language Learning, 6(1), 14-29.