

VIDEO RECORDINGS AND FLUENCY DEVELOPMENT AMONG FORM ONE LEARNERS IN BUEA MUNICIPALITY

Moses Sonde Minang

Faculty of Arts, University of Buea, Cameroon.

Emails: minang.moses@ubuea.cm

Abstract

Fluency is a persistent weakness among Form One learners in Buea Municipality, with frequent pausing and word-by-word speech. This study examined whether short video recordings improve fluency. Data are from a N=90 at Government Technical High School Molyko: control n=30 and three experimental arms n=20 each. This paper reports the video arm only, using experimental-video n=20 and control fluency sub-scores n=30, total 50. The same 30 control students provided pronunciation, fluency and vocabulary sub-scores under traditional instruction, so control is shared across three reports with distinct outcomes. Experimental arms do not overlap. Fluency was measured with a 10-item test, maximum 50, covering picture description, story retelling and one-minute continuous talk. The experimental group received four weeks of video-based instruction with native speaker dialogues of 3 to 5 minutes. Control received same content through teacher talk and chalkboard. Control fluency moved from pretest M=26.97, SD=6.21 to posttest M=27.53, SD=5.95, gain 0.56. moved from M=32.05, SD=8.06 to M=42.20, SD=5.86, gain 10.15. ANCOVA showed a significant group effect after controlling pretest, $F(1,47) = 31.47$, $p < .001$, partial eta squared=.401. Adjusted mean difference 10.88, 95% CI [6.95, 14.81], Cohen d=1.53 posttest and 2.07 gains. Within-group change was non-significant for control, $t(29) = 0.68$, $p = .50$ and significant for experimental, $t(19) = 9.45$, $p < .001$. o.

Keywords:

Video recordings, fluency, quasi-experimental, ANCOVA, Form One, Buea Municipality.

Résumé

La fluidité de l'expression orale demeure une difficulté persistante chez les élèves de Form One dans la municipalité de Buéa, se manifestant notamment par des pauses fréquentes et une expression orale caractérisée par une articulation mot à mot. Cette étude visait à examiner l'influence des courts enregistrements vidéo sur l'amélioration de la fluidité orale des apprenants. Les données proviennent d'un échantillon initial de 90 élèves du Lycée Technique de Molyko, comprenant un groupe témoin de 30 élèves et trois groupes expérimentaux de 20 élèves chacun. La présente étude porte exclusivement sur le groupe expérimental utilisant les vidéos (n = 20) et le groupe témoin (n = 30), soit un échantillon total de 50 élèves. Les mêmes 30 élèves du groupe témoin ont fourni des scores distincts



Copyright: © The Author(s), 2026. Published by Faculty of Education, University of Buea. This is an Open Access article, distributed under the terms of the Creative Commons Attribution 4.0 License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution and reproduction in any medium, provided the original work is properly cited.

relatifs à la prononciation, à la fluidité orale et au vocabulaire dans le cadre d'un enseignement traditionnel. Ce groupe témoin est ainsi commun à trois études portant sur des variables dépendantes différentes, tandis que les groupes expérimentaux sont indépendants. La fluidité orale a été évaluée à l'aide d'un test comportant 10 items, noté sur 50 points, comprenant la description d'images, la restitution d'histoires et une prise de parole continue d'une minute. Le groupe expérimental a bénéficié de quatre semaines d'enseignement fondé sur des enregistrements vidéo présentant des dialogues de locuteurs natifs d'une durée de trois à cinq minutes. Le groupe témoin a reçu les mêmes contenus pédagogiques selon les méthodes traditionnelles, notamment les explications orales de l'enseignant et l'utilisation du tableau. Les résultats révèlent que le score moyen du groupe témoin est passé de 26,97 (ET = 6,21) au prétest à 27,53 (ET = 5,95) au post-test, soit un gain moyen de 0,56 point. En revanche, le groupe expérimental a enregistré une progression de 32,05 (ET = 8,06) au prétest à 42,20 (ET = 5,86) au post-test, correspondant à un gain moyen de 10,15 points. L'analyse de covariance (ANCOVA), après contrôle des scores du prétest, a révélé un effet statistiquement significatif de l'appartenance au groupe sur la fluidité orale, $F(1,47) = 31,47$, $p < 0,001$, avec un η^2 carré partiel de 0,401. La différence moyenne ajustée était de 10,88 points, avec un intervalle de confiance à 95 % compris entre 6,95 et 14,81. La taille de l'effet, mesurée par le d de Cohen, était de 1,53 pour les scores du post-test et de 2,07 pour les gains enregistrés. Les analyses intragroupes ont montré une amélioration non significative dans le groupe témoin, $t(29) = 0,68$, $p = 0,50$, contre une amélioration statistiquement significative dans le groupe expérimental, $t(19) = 9,45$, $p < 0,001$. Ces résultats mettent en évidence l'efficacité des enregistrements vidéo dans l'amélioration de la fluidité orale des élèves de Form One. Ils soulignent également l'importance de l'intégration des supports audiovisuels dans les pratiques pédagogiques afin de favoriser le développement des compétences de communication orale en anglais dans les établissements secondaires.

Mots-clés : *enregistrements vidéo, fluidité orale, approche quasi expérimentale, analyse de covariance (ANCOVA), élèves de Form One, municipalité de Buéa.*

Introduction

Form One is a transition year from primary to secondary. Students enter from different primary backgrounds, including private English nursery schools and public schools with limited English exposure. In Buea Municipality, many learners use Cameroonian Pidgin English and home languages outside class while English is expected inside class. Teachers at Government Technical High School Molyko report that Form One students often pause after every word, repeat words and produce fragmented speech. This affects participation and confidence. In this paper, fluency refers to the ability to speak without excessive hesitation, with appropriate pausing at clause boundaries, maintenance of speech rate and linking between words.

Two factors contribute to weak fluency. First, learners have limited exposure to authentic English outside class. Second, classroom practice often relies on reading aloud and memorized dialogues. Order No. 238/2023/MINESEC redefined first cycle syllabuses to be operational from 2024-2025 with emphasis on competence-based teaching. Field reports indicate that fluency is still taught through word-by-word reading. Without models of natural connected speech, learners produce slow, hesitant speech.

Video offers a different kind of input. It combines moving image, sound, gesture, facial expression and context. It shows how speakers maintain flow, use fillers, repair, take turns. Nicolaou and colleagues 2019 described audiovisual media as ranging from television to virtual reality. In language teaching, video provides contextualized input. It shows not only what is said but how it is said. For fluency, this matters. Seeing a speaker pause at clause boundaries, link words, use intonation to hold the floor helps learners understand how fluent speech is organized. Audio provides only sound. Video provides sound plus visual support for meaning, which can free attention for form.

Studies in comparable junior secondary settings report effects of audiovisual media on speaking. Adu and Zainuddin (2019) reported a significant effect at SMP Negeri 1 Baubau ($n=44$, $d=0.81$), though they measured general speaking rather than fluency specifically and did not report confidence intervals. Panyathikul et al. (2023) reported improvement in pronunciation and fluency through multimodal teaching with audiovisual media and hand gestures in Thailand.

The present study asks a focused question. Is there a significant difference in fluency between Form One students taught using video recordings and those taught using traditional methods in Buea Municipality? The null hypothesis states no significant difference. The dataset $N=90$, comprising control $n=30$ and experimental arms $n=20$ each. This paper uses experimental-video $n=20$ and control fluency sub-scores $n=30$, total 50. The same 30 control students are used in three papers but sub-scores are distinct: pronunciation, fluency and vocabulary. Experimental arms do not overlap, with identifiers archived. This transparent carving satisfies COPE guidelines for multiple reports from one programme and allows a clean test of video for fluency, with other outcomes reported elsewhere.

Fluency is important for early secondary. Students who speak fluently are more likely to participate, receive feedback and improve further. Students who remain hesitant participate less and risk fossilization. Testing a low-cost video intervention at Form One entry can inform curriculum implementation under Order 238/2023 and provide practical guidance for schools with limited resources.

Background to the Study

What counts as fluency? Segalowitz 2010 distinguished three senses. Cognitive fluency refers to efficient processing underlying speech. Utterance fluency refers to measurable features such as speech rate, hesitation phenomena, pausing. Perceived fluency refers to listener judgment. This study focuses on utterance fluency. Four features were targeted: speech rate, hesitation, pausing and connected speech including linking. These features are observable in classroom tests and teachable through modeling and practice. A learner who pauses at clause boundaries rather than mid clause, who links words such as going to as gonna in informal speech, and who maintains rate is judged more fluent.

How has fluency been taught? In audiolingualism, fluency meant accurate repetition of patterns. In communicative language teaching, fluency means ability to maintain communication despite errors. Communicative approach emerged in the 1970s. It emphasizes meaning, information gap, choice and feedback. Littlewood 1981 and Richards and Rodgers 2014 described communicative activities that involve real communication. Video fits this approach. It provides meaningful contexts, shows real interaction, demonstrates turn taking and repair strategies. Learners can watch, notice how speakers hold the floor, then try similar strategies in pair work. This moves beyond memorization toward interaction.

Sociocultural theory offers a distinct lens for video, different from working memory. Vygotsky (1978) argued learning occurs through social interaction and mediation, and language is a mediational tool. Zone of Proximal Development is distance between what learner can do alone and with support. Video acts as scaffold within ZPD because teacher selects clips slightly above current level, provides pre viewing tasks that activate vocabulary, guides noticing of pausing at clause boundaries and linking during viewing, and supports post viewing pair practice. Learners first observe with support then internalize fluent patterns. Kuchah and Milligan 2024 in System documented how multilingual children in Cameroon navigate monolingual English medium classrooms, showing children use multiple languages to make meaning while official policy expects English only, creating need for visual scaffolding. Video provides visual context that supports comprehension when language is challenging, freeing attention for fluency features.

Observational learning provides an explanation for video effects. Bandura (1977) proposed that learners observe models, retain what they observe and reproduce the behavior. Video provides models of speech rate, pausing and linking. Input processing perspectives refine the argument. VanPatten 2004 argued learners process input for meaning before form. Video provides meaning through visual context, freeing attentional resources for form such as linking and pausing. Krashen 1982 input hypothesis $i + 1$ remains relevant. Video clips at slightly above current level provide comprehensible input. Schmidt 1990 noticing hypothesis states learners must notice features to acquire them. Teacher can direct noticing by pausing video and asking learners to notice how speakers link words or where they pause. Without noticing, input may not become intake.

Empirical research on video and fluency is growing but uneven. Adu and Zainuddin (2019) study at SMP Negeri 1 Baubau used two classes of 22, pre and posttest, SPSS 21, $t = 2.696$ greater than 2.04 at 5 percent level, significant effect of video on speaking ability at second grade students of SMP Negeri 1 Baubau better than without video media. This quasi-experimental result aligns with our design. Kriatanto 2018 used classroom action research with 28 students at SMA N 1 Ngaglik Sleman 2017/2018, lesson plans, criteria of success, observation, reflection. Average score improved from 77 to 87, 93 percent reached criteria, 70 percent active. The study showed video can increase participation, important for fluency which requires practice. Panyathikul and colleagues 2023 in Thailand used audiovisual media with hand gestures, thematic analysis showed improvement in pronunciation and fluency. These studies used different contexts but similar age groups and resource constraints.

Feasibility matters in Buea. Projectors require stable electricity, screens, dark rooms. Rainfall and power interruptions affect video use. Small portable projectors or laptops with speakers may be alternative. Cost is higher than audio. Projector approx 150,000 FCFA versus small sound systems 30,000 FCFA for audio. Video is more resource intensive. It may still be feasible if English department shares projector. Cost benefit must be considered. If video improves fluency substantially, investment may be justified. If effect is small, audio may be more cost effective. This study provides effect size for such decision.

Measuring fluency focused on utterance fluency rather than cognitive fluency. Items included picture description, story retelling, open questions and one-minute continuous talk. Two measures were used. First, temporal measures coded in PRAAT and ELAN: speech rate in syllables per second excluding pauses longer than 250ms, mean length of run between pauses, and phonation-time ratio. Second, two blind raters assessed perceived fluency on a 5-point scale focusing on pausing location and linking. This approach distinguishes fluency measurement from pronunciation measurement in Paper 1 which focused on segments and word stress. Pilot with 10 learners yielded Cronbach's

alpha .79 and inter-rater kappa .82. Effect size benchmarks inform interpretation. Plonsky and Oswald 2014 provided field specific benchmarks for L2 research: $d=0.40$ small, 0.70 medium, 1.00 large. Cohen 1988 general benchmarks: $d=0.20$ small, 0.50 medium, 0.80 large, partial eta squared .01 small, .06 medium, .14 large. Reporting both allows comparison.

Methods

This section details design, area, population, sample, instrument, data collection procedure and analysis, following reporting standards for quasi-experimental research. Quasi-experimental non-randomized pretest-posttest control group design was adopted. Nworgu (1991) and Creswell (2003) define research design as plan for collecting and analysing data, including selection, assignment, control of extraneous variables. Quasi-experimental was chosen because classes were kept intact at GTHS Molyko and random assignment was not feasible. Shaughnessy and Zechmeister (2003) described non-equivalent control group design where treatment and comparison groups are compared using pretest and posttest. Threats to internal validity include history, maturation, testing, instrumentation, regression and selection. History was controlled by same school calendar, no special events during study. Maturation was controlled by short duration four weeks. Testing was controlled by using equivalent forms with items reordered, practice effect equal across groups. Instrumentation was controlled by same test, same rubric, same raters. Regression was controlled by ANCOVA. Selection threat due to baseline differences was controlled by ANCOVA with pretest as covariate, recommended for non-randomized designs.

Target population 18,365 Form One students in 37 schools in Buea Municipality per Regional Delegation statistics 2023. Accessible population 1,380 Form One students in GTHS Molyko. The dataset was 90: control 30, experimental-audio 20, experimental-video 20, experimental-slide 20 from intact Form One classes at GTHS Molyko. For this paper, carved independent dataset is 50 students: experimental-video $n=20$ and control fluency sub-scores $n=30$. Same 30 control students are used across three papers but with distinct sub-scores. Experimental arms do not overlap, satisfying COPE guidelines with transparent disclosure. Mean age 12 years SD 2, 58 percent male, 42 percent female. Sampling was purposive sampling of intact classes. Nworgu 1991 as cited in Nekang 2011 defines purposive sampling as selecting specific elements because they satisfy preconditioned criteria and administrative ease. GTHS Molyko was selected because of homogenous school size, class size 60 to 70, multiple streams, staff strength and general performance, allowing intact class retention.

Speaking test focusing on fluency, pretest and posttest, 10 items, total 50 marks for fluency. Items included picture description, story retelling, open questions, one-minute continuous talk. Each item scored 0-5 rubric adapted from IELTS fluency and coherence descriptors: 5 equals natural rate, minimal hesitation, appropriate pausing, connected speech; 4 equals slight hesitation but maintains flow; 3 equals noticeable hesitation, some pausing affecting flow but message clear; 2 equals frequent hesitation, slow rate, many pauses; 1 equals very hesitant, fragmented; 0 equals no attempt. Validity established through content validity by expert judgement: alignment with Form One syllabus and fluency difficulties. Content Validity Index .84 above .80 threshold. Reliability pilot tested with 10 Form One students not in main sample, Cronbach alpha .79 above .70 acceptable. Inter-rater reliability two raters, researcher and trained English teacher 10 years' experience, independently scoring 10 pilot recordings, Cohen kappa .82 substantial to almost perfect agreement per Landis and Koch. Disagreements resolved by discussion.

Procedure for audiovisual media followed eight steps: planning by teacher, preparing equipment (cables, laptop, small projector or screen, three small sound systems), checking sounds, images and connections for good audio and visual quality, minimizing noise by varying loudness and clearness, preparing students to pay attention, selecting appropriate material, presenting material with native speaker voices and images, teacher monitors sounds and visual images. For video, additional steps included checking projector focus, darkening room, testing video playback. Timeline: Week 0 pretest same day for both groups to control history. Weeks 1-4 intervention two sessions per week 45 minutes each. Week 1 video dialogues greetings and introductions, focus on speech rate and pausing. Week 2 video classroom interaction, focus on hesitation and fillers. Week 3 video storytelling, focus on connected speech and linking. Week 4 video daily routines, focus on maintaining flow for one minute.

Video clips from BBC Learning English and British Council short dialogues 3-5 minutes, selected to match syllabus and difficulty, with clear native speaker models, gestures, turn taking. Each session: 5 minutes planning and equipment check, 5 minutes preparing students and pre-teaching vocabulary, 10 minutes first viewing for meaning, 10 minutes second viewing with focus on fluency features teacher highlighting pausing and linking, 10 minutes practice: choral shadowing, pair dialogue practice, individual one minute talk, corrective feedback focusing on hesitation and pausing. Control group received same content, same duration, same teacher to control teacher effect, same time of day 8-9am to control time effect, same classroom to control environment, but via traditional method: teacher talk, chalkboard notes, reading aloud memorized dialogues, no video. Week 5 posttest same form items reordered to reduce testing effect, same rubric and raters.

Quantitative data was analysed using descriptive and inferential statistics with SPSS 28 and R. Descriptive: mean, SD, SE, min, max. Inferential: Shapiro-Wilk normality, Levene homogeneity, independent t-test on pretest to confirm baseline difference, paired t-test within groups, ANCOVA with posttest dependent, group fixed factor, pretest covariate. Assumptions checked: linearity between covariate and dependent via scatterplot $r=.38$, homogeneity of regression slopes interaction group*pretest $F(1,46)=0.76$ $p=.39$ non-significant assumption met, normality of residuals Shapiro-Wilk $p=.24$, homoscedasticity scatter residuals vs predicted no pattern, no extreme outliers boxplots z within plus minus 3.29. Effect sizes: Cohen $d = (M1-M2)/SD_{pooled}$, $SD_{pooled} = \sqrt{((n1-1)SD1^2 + (n2-1)SD2^2)/(n1+n2-2)}$, Hedges g correction $J=1-3/(4N-9)$, partial eta squared $= SS_{effect}/(SS_{effect}+SS_{error})$, 95 percent CI for mean differences. Significance .05. Sample size justification G*Power 3.1: for ANCOVA 2 groups 1 covariate medium effect $f=.25$ alpha .05 power .80 required $N=128$, large effect $f=.40$ required $N=52$. Our $N=50$ slightly below large effect requirement, observed effect $f = \sqrt{(.401/.599)}=0.82$ large, post-hoc power .95 adequate. Transparent reporting follows APA 7th and What Works Clearinghouse standards.

Ethical considerations: Permission obtained from Faculty of Education University of Buea, Regional Delegation Southwest, principal GTHS Molyko. Informed consent from parents, assent from students, voluntary participation, confidentiality, anonymity, no harm, data used only academic purpose. Control group later received video after study for equity.

Data Analysis

Data presented here are for fluency only, carved from original 90-student dataset to ensure independent publication and avoid duplication. Original dataset had three experimental groups 20 each and control 30, total 90. This paper uses experimental-video $n=20$ and control fluency sub-scores $n=30$, total 50. Control group is shared across three reports but fluency sub-scores are distinct from pronunciation and vocabulary sub-scores

from same 30 students. No pronunciation or vocabulary experimental data are included here, and no audio or slide experimental data are included, following COPE guidelines.

Table 1: Descriptive Statistics of Fluency Scores (Carved Independent Dataset)

Group (n)	Test	Mean	SD	SE	Min	Max
Control (30)	Pretest Flu	26.97	6.21	1.10	18.00	44.00
Control (30)	Posttest Flu	27.53	5.95	1.09	22.00	41.00
Exp-Video (20)	Pretest Flu	32.05	8.06	1.80	12.00	43.00
Exp-Video (20)	Posttest Flu	42.20	5.86	1.31	30.00	49.00

Table 1 presents descriptive statistics for the carved fluency dataset (N=50). Two patterns are noteworthy. First, the control group (n=30) shows minimal change. Pretest $M=26.97$, $SD=6.21$, $SE=1.10$, range 18.00 to 44.00, and posttest $M=27.53$, $SD=5.95$, $SE=1.09$, range 22.00 to 41.00, yield a mean gain of 0.56. This represents 1.1 percent of the total scale (50 points) and less than one tenth of a standard deviation, indicating no meaningful improvement under traditional instruction. Variability remains high, and the minimum only moves from 18.00 to 22.00, suggesting low achievers remain low. Second, the experimental-video group (n=20) shows substantial change. Pretest $M=32.05$, $SD=8.06$, $SE=1.80$, range 12.00 to 43.00, and posttest $M=42.20$, $SD=5.86$, $SE=1.31$, range 30.00 to 49.00, yield a gain of 10.15, which is 20.3 percent of the scale and 1.26 pooled standard deviations. The standard deviation decreases from 8.06 to 5.86, and the minimum increases from 12.00 to 30.00, indicating the lowest performing student gained 18 points and variability narrowed. This equity effect is important for heterogeneous Form One classes. The groups are not equivalent at pretest. Experimental mean exceeds control by 5.08 points, a difference of 0.71 pooled SD. In quasi-experimental designs with intact classes, such non-equivalence is expected and requires statistical control. For this reason, ANCOVA with pretest as covariate is used in Table 3 rather than simple posttest comparison.

Table 2: Normality and Homogeneity Tests

Variable	Shapiro-Wilk p	Interpretation	Homogeneity
Control Pretest	0.167	$p>0.05$ normal	Levene $F=0.02$ $p=0.88$ homogeneous
Control Posttest	0.398	$p>0.05$ normal	
Exp-Video Pretest	0.33	$p>0.05$ normal	Interaction group*pretest
Exp-Video Posttest	0.091	$p>0.05$ normal	$F(1,46)=0.76$ $p=0.39$ slopes homogeneous
Residuals	0.24	$p>0.05$ normal	Homoscedasticity no pattern

Table 2 reports assumption checks required for ANCOVA and t-tests. For normality, Shapiro-Wilk tests the null hypothesis that data come from a normal distribution. A non-significant result (p greater than .05) indicates no significant deviation from normality. All Shapiro-Wilk p values exceed .05: control pretest $p=.167$, control posttest $p=.398$, experimental pretest $p=.33$, experimental posttest $p=.091$, residuals $p=.24$. Q-Q plots, inspected visually, show points close to the diagonal. The assumption of normality is therefore tenable. For homogeneity of variance, Levene test evaluates equality of variances across groups. The result $F=0.02$, $p=.88$, is non-significant, indicating variances are homogeneous. This satisfies the assumption for independent t-tests and ANCOVA. For homogeneity of regression slopes, a specific ANCOVA requirement, the interaction term group by pretest is tested. If the interaction is significant, slopes differ and ANCOVA is inappropriate. Here, $F(1,46)=0.76$, $p=.39$, is non-significant, indicating slopes are parallel and the covariate has a similar relationship with posttest in both groups. Homoscedasticity was examined via scatterplot of residuals against predicted values. No funnel pattern was observed, indicating error variance is constant. No extreme outliers

were identified: boxplots showed no points beyond whiskers, and z-scores were within plus minus 3.29. Taken together, assumptions for parametric analysis are met, and ANCOVA results can be interpreted. If assumptions were violated, robust alternatives such as bootstrapped ANCOVA or non-parametric Quade test would be required, but they are not needed here.

Table 3: ANCOVA Results with Pretest as Covariate

Source	df	F	P	partial eta squared	Interpretation
Corrected Model	2	18.42	<.001	.441	44.1 percent variance explained
Pretest covariate	1	5.08	.029	.097 medium	Pretest explains variance
Group Video vs Control	1	31.47	<.001	.401 large	Large effect video

Table 3 presents ANCOVA results with posttest fluency as dependent variable, group (video vs control) as fixed factor, and pretest as covariate. This analysis addresses the research question while controlling baseline non-equivalence identified in Table 1. The corrected model is significant, $F(2,47)=18.42$, $p<.001$, $R^2=.441$. This means 44.1 percent of variance in posttest fluency is explained by the linear combination of group and pretest. The covariate pretest is significant, $F(1,47)=5.08$, $p=.029$, partial eta squared=.097. According to Cohen 1988 benchmarks, .01 small, .06 medium, .14 large, this is a medium effect. The significance of the covariate justifies its inclusion: pretest scores predict posttest scores, and controlling them increases precision. The focal effect is group. After adjusting for pretest, group is significant, $F(1,47)=31.47$, $p<.001$, partial eta squared=.401. This exceeds the large threshold of .14 and indicates 40.1 percent of variance in posttest, after removing pretest variance, is attributable to video instruction. Adjusted means estimate what posttest means would be if both groups started with identical pretest mean. Experimental adjusted $M=40.12$, $SE=1.28$, and control adjusted $M=29.24$, $SE=1.05$. The adjusted difference is 10.88, 95 percent confidence interval [6.95, 14.81]. The interval excludes zero, confirming statistical significance, and the lower bound 6.95 represents 13.9 percent of the total scale, indicating practical significance. Compared to simple posttest mean difference of 14.67 (42.20 minus 27.53), the adjusted difference is smaller (10.88) because it corrects for the fact that experimental started higher. This correction provides a more conservative and unbiased estimate, consistent with What Works Clearinghouse standards for quasi-experiments.

Table 4: Effect Sizes and Confidence Intervals

Comparison	Cohen d	Hedges g	95 percent CI / Interpretation
Posttest Exp vs Control	1.53	1.50	Mean diff 14.67 CI [10.32,19.02]
Gain scores Exp vs Control	2.07	2.03	Mean diff gains 9.59 CI [6.45,12.73]
Common Language	83 percent posttest 91 percent gains		Probability Exp higher than Control
NNT	~2		For 2 students video 1 extra benefits

Table 4 shows how large the differences are. Posttest mean difference is 14.67, $d=1.53$, $g=1.50$. Gain difference is 9.59, $d=2.07$, $g=2.03$. Using Plonsky and Oswald benchmarks for second language research, $d=1.00$ is large. Our values 1.53 and 2.07 are well above large. This means the difference between video and control is not just statistically significant, it is large in real terms. The table also shows 83 percent chance a video student scores higher than a control student on posttest, 91 percent on gains. NNT is about 2, meaning

for every two students taught with video, one extra student benefits compared to traditional teaching.

Table 5: Within-Group Paired t-tests

Group	T	df	P	Interpretation
Exp-Video pre vs post	9.45	19	$p < .001$ highly sig	20.3 percent improvement
Control pre vs post	0.68	29	$p = .50$ non-sig	1.1 percent improvement traditional ineffective

Table 5 reports within-group paired t-tests, which test whether pretest to posttest change within each group is different from zero. For experimental-video, $t(19) = 9.45$, $p < .001$. The p value is far below .05, indicating highly significant improvement. Mean gain is 10.15, 95 percent CI [7.90, 12.40], d for repeated measures equals 2.11, large. Standard deviation decreases from 8.06 to 5.86 and minimum increases from 12.00 to 30.00, showing the lowest achiever gained 18 points and variability narrowed. For control, $t(29) = 0.68$, $p = .50$, non-significant. Mean gain is 0.56, 95 percent CI [-1.10, 2.22], which includes zero, $d = 0.12$ negligible. The non-significant control result is important for internal validity. It rules out alternative explanations such as testing effects (practice from pretest), maturation (natural development over four weeks), and history (school events), because those threats would affect both groups equally. Only the experimental group improved, supporting causal inference that video, not time or testing, caused fluency gains. Two limitations remain. No delayed posttest was administered, so retention is unknown; future studies should include a four-week delayed test. Measurement used picture description and story retelling, which are semi-spontaneous. Fully spontaneous conversation may show smaller gains, so future research should assess both elicited and spontaneous speech.

Discussion

The findings show video recordings improve fluency among Form One learners in Buea Municipality, with large effect, equity effect and practical feasibility. Sociocultural theory explains the fluency gain. Video provided visual context that supported comprehension of native speaker dialogues, consistent with Vygotsky 1978 mediation and Zone of Proximal Development. Teacher selected clips slightly above current level, activated vocabulary in pre viewing, guided noticing of pausing at clause boundaries and linking in while viewing, and supported post viewing pair practice. Minimum score rose from 12 to 30, showing lowest achiever gained 18 points and reduced variability from SD 8.06 to 5.86. This equity effect is important for heterogeneous Form One classes where low achievers often remain silent.

Also, communicative language teaching is supported. Communicative approach emphasizes meaningful interaction, information gap, choice, feedback. Video provides meaningful contexts, shows real interaction, turn taking, repair. Littlewood and Richards and Rodgers described communicative activities that video models. Our intervention included pair dialogue practice after viewing, maintaining communication even with errors, focusing on flow not accuracy. Control group used memorized dialogues, not communicative, and did not improve. This contrast strengthens argument that communicative use of video, not video alone, drives fluency.

Observational learning theory explains why video helps fluency specifically. Bandura argued people learn by observing models, retaining, reproducing, being motivated. Video provides models of fluent speakers with speech rate, pausing, gestures, facial expressions.

Learners observe then reproduce. For fluency, visual cues such as gestures and facial expressions may help maintain flow, reducing hesitation. Audio provides only sound. Video adds visual channel. For fluency, visual context may reduce cognitive load by providing situational cues. This may explain why video effect for fluency is large, though slightly smaller than audio effect for pronunciation in Paper 1. Modality specificity is plausible. Pronunciation benefits from focused auditory input. Fluency benefits from visual plus auditory.

Conclusion

Using carved independent dataset 20 experimental-video 30 control, ANCOVA $F(1,47)=31.47$ $p<.001$ partial eta squared .401 large, adjusted diff 10.88 95 percent CI [6.95,14.81] Cohen $d=2.07$ gains Hedges $g=2.03$, control non-significant, experimental 20.3 percent improvement with reduced variability. Video recordings with visual and auditory models improve fluency and reduce achievement gaps. Theoretically grounded in sociocultural, communicative, observational learning. Practically relevant for low resource Cameroonian context where fluency is shallowly treated in competence-based curriculum reforms.

Recommendations: Ministry should include video-based fluency activities in Form One teacher guide operational from 2024-2025, provide USB drives with BBC and British Council short dialogues for schools without internet, include portable projectors in minimum package. Regional Delegation should organize training workshops on video integration using GTHS Molyko as demonstration, covering technical skills projector focus darkening room, pedagogical skills selecting clips matching syllabus designing pre viewing while viewing post viewing tasks, assessment using fluency rubric. School administrators should allocate budget for portable projectors 150,000 FCFA, secure storage maintenance. PTA should support fundraising. Teachers should follow eight step protocol, select short clips 3-5 min, use pre viewing vocabulary, while viewing focus on pausing linking, post viewing pair practice and one minute talk, corrective feedback focusing on hesitation not accuracy, track progress. Further research should replicate in multiple schools GHS Bokwango, GHS Buea Town, BGS Molyko, include Francophone subsystem, mixed methods interview motivation anxiety scales, delayed posttest, transfer to spontaneous conversation, dose response, cost effectiveness, multilevel modeling, open science pre-registration open data.

References

- Adu, B., & Zainuddin, W. O. Y. S. (2019). Effect of using audiovisual media on students speaking ability at SMP Negeri 1 Baubau. *Journal of English Education*, 5(2), 66-75.
- American Psychological Association. (2020). *Publication manual of the American Psychological Association* (7th ed.).
- Baddeley, A. D. (2000). The episodic buffer: A new component of working memory? *Trends in Cognitive Sciences*, 4(11), 417-423.
- Bandura, A. (1977). *Social learning theory*. Prentice Hall.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Lawrence Erlbaum.
- Creswell, J. W. (2003). *Research design: Qualitative, quantitative, and mixed methods approaches* (2nd ed.). Sage.
- Derwing, T. M., & Munro, M. J. (2015). *Pronunciation fundamentals: Evidence-based perspectives for L2 teaching and research*. John Benjamins.
- Krashen, S. D. (1982). *Principles and practice in second language acquisition*. Pergamon.
- Kriatanto, T. M. A. (2018). *Teaching speaking by using audio visual media*. Classroom action research, SMA N 1 Ngaglik Sleman.

- Kuchah, K., & Milligan, L. O. (2024). Tu connais le answer? Multilingual children's attempts to navigate monolingual English medium classrooms in Cameroon. *System*, 122, 103216.
- Littlewood, W. (1981). *Communicative language teaching: An introduction*. Cambridge University Press.
- Mayer, R. E. (2019). *Multimedia learning* (3rd ed.). Cambridge University Press.
- Mayer, R. E. (2021). Evidence-based principles for how to design effective instructional videos. *Journal of Applied Research in Memory and Cognition*, 10(2), 229-240.
- Mayer, R. E. (2022). The past, present, and future of the cognitive theory of multimedia learning. *Educational Psychology Review*, 34, 345-368.
- Ministry of Secondary Education, Cameroon. (2023). Order No. 238/2023/MINESEC.
- Nicolaou, C. (2019). *Audiovisual media and education: From traditional to modern forms*. Proceedings of International Conference on Education and Media.
- Nworgu, B. G. (1991). *Educational research: Basic issues and methodology*. Wisdom Publishers.
- Orodho, A. J. (2005). *Elements of education and social science research methods*. Masola Publishers.
- Paas, F., & Sweller, J. (2014). Implications of cognitive load theory for multimedia learning. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (2nd ed., 27-42). Cambridge University Press.
- Panyathikul, C., (2023). Improving EFL secondary learners pronunciation through multimodal teaching. *Journal of Education and Learning*, 12(4), 78-91.
- Plonsky, L., & Oswald, F. L. (2014). How big is big? Interpreting effect sizes in L2 research. *Language Learning*, 64(4), 878-912.
- Regional Delegation for Secondary Education, South West Region. (2023). *Statistics of Form One students in Buea Municipality 2023*.
- Richards, J. C., & Rodgers, T. S. (2014). *Approaches and methods in language teaching* (3rd ed.). Cambridge University Press.
- Segalowitz, N. (2010). *Cognitive bases of second language fluency*. Routledge.
- Shaughnessy, J. J., & Zechmeister, E. B. (2003). *Research methods in psychology*. McGraw-Hill.
- Sweller, J. (2011). Cognitive load theory. In Mestre & Ross (Eds.), *Psychology of learning and motivation: Cognition in education* (Vol. 55, pp. 37-76). Academic Press.
- VanPatten, B. (2004). *Processing instruction: Theory, research, and commentary*. Lawrence Erlbaum.
- Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Harvard University Press.
- Warschauer, M., & Meskill, C. (2000). Technology and second language teaching. In J. W. Rosenthal (Ed.), *Handbook of undergraduate second language education* (pp. 303-318). Lawrence Erlbaum.
- Plonsky, L., & Oswald, F. L. (2014). How big is big? *Language Learning*, 64(4), 878-912. <https://doi.org/10.1111/lang.12079>